Electronic Transactions on Numerical Analysis.
Volume 39, pp. 75-101, 2012.

ISSN 1068-9613.

THE COMPLETE STAGNATION OF GMRES FOR $N \leq 4$*

GÉRARD MEURANT†

Abstract. We study the problem of complete stagnation of the generalized minimum residual method for real matrices of order $n \leq 4$ when solving nonsymmetric linear systems $Ax = b$. We give necessary and sufficient conditions for the non-existence of a real right-hand side b such that the iterates are $x^k = 0$, $k = 0, \dots, n-1$, and $x^n = x$. We illustrate these conditions with numerical experiments. We also give a sufficient condition for the non-existence of complete stagnation for a matrix A of any order n.

Key words. GMRES, stagnation, linear systems

AMS subject classifications. 15A06, 65F10

1. Introduction. We consider solving a linear system

$$Ax = b, \tag{1.1}$$

where A is a nonsingular real matrix of order n with the generalized minimum residual method (GMRES), which is a Krylov method based on the Arnoldi orthogonalization process; see Saad and Schultz [46]. The initial residual is denoted as $r^0 = b - Ax^0$ where x^0 is the starting vector. Without loss of generality we will choose $x^0 = 0$, which gives $r^0 = b$, and we assume $\|b\| = 1$, where $\| \cdot \|$ is the l_2 norm. The Krylov subspace of order k based on A and r^0, denoted as $\mathcal{K}_k(A, r^0)$, is $span\{r^0, Ar^0, \dots, A^{k-1}r^0\}$. The approximate solution x^k at iteration k is sought as $x^k \in x^0 + \mathcal{K}_k(A, r^0)$ such that the norm of the residual vector $r^k = b - Ax^k$ is minimized.

Complete stagnation of GMRES corresponds to $\|r^k\| = \|b\|$, $k = 0, \dots, n-1$, and $\|r^n\| = 0$. Since $\|r^{n-1}\| \neq 0$ implies that the degree of the minimal polynomial of A is equal to n, we assume that the matrix A is non-derogatory. This means that, up to the sign, the characteristic polynomial is the same as the minimal polynomial. We are interested in characterizing the real right-hand sides b which give complete stagnation for a given matrix A. We call those b stagnation vectors. We will give necessary and sufficient conditions for the non-existence of such vectors b if $n \leq 4$ and only sufficient conditions for $n > 4$. The problem of GMRES complete stagnation was considered by Zavorin, O'Leary and Elman [60] assuming that the matrix A is diagonalizable; see also [59]. Sufficient conditions for non-stagnation in particular cases were given in [47, 48]. We have the well-known general characterization of complete stagnation that is also valid when the matrix A is complex; see, for instance, [33] or [60].

THEOREM 1.1. *We have complete stagnation of GMRES if and only if the right-hand side b of the linear system (1.1) satisfies*

$$(b, A^j b) = 0, \quad j = 1, \dots, n-1. \tag{1.2}$$

The inner product in $\mathbb{C}^n$ is defined as $(x, y) = \sum_{i=1}^n x_i \bar{y}_i$, where the bar denotes the complex conjugate. The characterization of Theorem 1.1 shows that complete stagnation is not possible if 0 is outside the field of values of any of the matrices A^j, $j = 1, \dots, n-1$, which are powers of A. This is the case if the symmetric part of any of these matrices is definite. The field of values of a matrix B is defined as

$$W(B) = \{(Bx, x),\ x \in \mathbb{C}^n, \|x\| = 1\}.$$

*Received June 10, 2011. Accepted February 15, 2012. Published online on May 2, 2012. Recommended by H. Sadok.

†(gerard.meurant@gmail.com).

ETNA
Kent State University
http://etna.math.kent.edu

We add the condition $\|b\| = 1$ to (1.2) since if b is a solution, then αb is also a solution of (1.2). Note that if b is a solution, then $-b$ is also a solution. So the number of stagnation vectors is even. We now have n nonlinear equations in n unknowns (the components of b), and the question is to know if there exists at least a real solution and, eventually, how many. Since we are interested in real vectors b, the system defined by (1.2) and $b^T b = 1$ is a polynomial system. Study of the solution set of polynomial systems is the domain of algebraic geometry. However, most of the problems that are considered in the literature have integer or rational coefficients. For polynomial systems with real coefficients; see Stetter [50].

The content of this paper is as follows. Section 2 considers the problem of existence of solutions to the complete stagnation problem. For a general order n we give a sufficient condition for the non-existence of stagnation vectors and we prove that the number of real stagnation vectors (which is between 0 and 2^n) is a multiple of 4. Section 3 gives necessary and sufficient conditions for $n = 3$ and $n = 4$. These conditions are based on known results about the simultaneous annealing of several quadratic forms defined by symmetric matrices. Existence or non-existence of stagnation vectors are illustrated by numerical examples in Section 4. Finally, Section 5 provides some conclusions and perspectives.

Throughout the paper $\imath$ denotes $\sqrt{-1}$. For our application, that is, the study of GMRES stagnation, the matrix A_i will denote $A^i + (A^i)^T$ for integer values of i.

2. Existence of solutions. We have already seen a necessary condition for the existence of solutions in Theorem 1.1 that can be rephrased as follows.

THEOREM 2.1. *A necessary condition to have a stagnation vector is that* 0 *is in the field of values of* A^j, $j = 1, \dots, n-1$.

Proof. Clearly a solution b has to be in the intersection of the inverse images of 0 for the functions $b \mapsto (A^j b, b)$, $j = 1, \dots, n-1$. If at least one of these sets is empty, there is no solution to the nonlinear system. □

The converse of this theorem is false for $n \geq 3$ and real vectors b. We will give counter-examples in Section 4. There exist matrices A with 0 in the fields of values of A^j for $j = 1, \dots, n-1$, and no real stagnation vector. The nonlinear system (1.2) can be transformed into a problem with symmetric matrices since with a real matrix B and a real vector b, we have the equivalence

$$b^T B b = 0 \Leftrightarrow b^T (B + B^T) b = 0.$$

Therefore, when A is real and if we are looking for real vectors b, we can consider the polynomial system

$$b^T (A^j + (A^j)^T) b = b^T A_j b = 0, \quad j = 1, \dots, n-1, \quad b^T b = 1, \tag{2.1}$$

with symmetric matrices A_j. The polynomial system (2.1) corresponds to the simultaneous annealing of $n-1$ quadratic forms defined by symmetric matrices and a nonzero vector. We are only interested in real solutions of (2.1) since complex ones do not provide a solution of the stagnation problem. The following straightforward theorem gives a sufficient condition for the non-existence of real stagnation vectors.

THEOREM 2.2. *Let* A *be a real matrix of order* n. *A sufficient condition for the non-existence of unit real stagnation vectors* b *is that there exist a vector* μ *with real components* μ_j, $j = 1, \dots, n-1$, *such that the matrix*

$$A(\mu) = \sum_{j=1}^{n-1} \mu_j A_j$$

is (positive or negative) definite.

ETNA
Kent State University
http://etna.math.kent.edu

Proof. Let us assume that the matrix $A(\mu)$ is definite for a given choice of coefficients. If there is a real b satisfying equation (2.1), multiply $A(\mu)$ by b from the right and b^T from the left. Then

$$b^T A(\mu) b = b^T (\sum_{j=1}^{n-1} \mu_j A_j) b = \sum_{j=1}^{n-1} \mu_j b^T A_j b = 0,$$

but since $A(\mu)$ is definite, this gives $b = 0$ which is impossible since $b^T b = 1$. ∎

Therefore, to have at least one real stagnation vector, the matrix $A(\mu)$ must be indefinite for any choice of the real numbers μ_j. Of course, as we already know, there is no stagnation vector if any of the matrices A_j is definite. The converse of Theorem 2.2 may not be true in general and it would be interesting to find counter-examples. However, as we will see in the next section, the converse is true for $n \leq 4$. Therefore, to find counter-examples one has to consider matrices of order $n \geq 5$. Moreover, we do not deal with any number of quadratic forms. For matrices of order n, we have exactly $n - 1$ quadratic forms. Necessary and sufficient conditions for the existence of stagnation vectors obtained with different techniques will be given in a forthcoming paper [39].

For the system (2.1) we are interested in the existence and the number of solutions of polynomial systems. There is an extensive literature on this topic. One can use, for instance, reference [22] where we have the following results that were obtained using homotopy. They show that, generically, there exist solutions.

THEOREM 2.3 (Theorem 2.1 of Garcia and Li [22]). *Let w represent the coefficients of the polynomial system $P(x, w) = 0$ of n equations in n unknowns and let d_i be the total degree of equation i. Then for all w except in a set of measure zero, the system has exactly $d = \prod_{i=1}^{n} d_i$ distinct solutions.*

For our problem, the degree of each equation is $d_i = 2$, and for A and b real we have n equations. Hence the maximum number of solutions is $d = 2^n$ as it is well-known. However, this result is not completely satisfactory since the vector b can be such that the coefficients are in the set of measure zero.

There is a more precise statement in [22]. Let H be the highest order system related to P. The system H is obtained from P by retaining only the terms of degree d_i in equation i of P.

THEOREM 2.4 (Theorem 3.1 of Garcia and Li [22]). *If $H(z) = 0$ has only the trivial solution $z = 0$, then $P(x) = 0$ has $d = \prod_{i=1}^{n} d_i$ solutions.*

For our system (2.1) and for real b, H is the same as P except for the constant term in the equation $b^T b - 1 = 0$ since all the other terms are of degree 2. It is clear that this system cannot have a solution different from zero. Hence, the system (2.1) has exactly 2^n solutions. The number 2^n is known as the Bézout number, named after the French mathematician Etienne Bézout (1730–1783). However, not much seems to be known about the fact that the solutions are real or complex. Unfortunately, the complex solutions of the system (2.1) are not solutions of the stagnation problem. There are ways to count the number of real solutions in the literature, but they are almost as complicated as computing all the solutions. However, we have the following result about the number of real solutions.

THEOREM 2.5. *Let A be a real matrix of order $n \geq 2$. The number of real solutions of the polynomial system (2.1) is a multiple of 4.*

Proof. The total number of solutions is 2^n. If b is a complex solution of (2.1), then also $-b, conj(b)$, and $-conj(b)$ are solutions where $conj(b)$ is a vector with the complex conjugates of the elements of b. Note that we have four different solutions unless b is purely imaginary, $b = \imath c$, $c \in \mathbb{R}^n$, because then $conj(b) = -b$. But such a vector b cannot be a solution since the last equation will be $b^T b = \imath^2 c^T c = 1$ and $\|c\|^2 = -1$ which is impossible.

Hence, the number of complex solutions is a multiple of 4, say $4m$. The number of real solutions is $2^n - 4m = 4(2^{n-2} - m)$ for $n \geq 2$. This shows that the number of real solutions is a multiple of 4. ∎

3. The case $n \leq 4$. We start this section by recalling known results about quadratic forms and then we use those results for characterizing non-stagnation for $n \leq 4$. The reason we are restricted to $n \leq 4$ is that the results in the literature correspond to two or three quadratic forms and this only yields results for $n = 3$ or $n = 4$ for the stagnation problem.

3.1. Results on quadratic forms. The simultaneous annealing of two quadratic forms defined by symmetric matrices of order n has been studied for a long time. The story of solutions to this problem seems to begin with Paul Finsler in 1937 [21] with what is now known as Finsler's theorem or Debreu's lemma. Using Uhlig's notation [54, 55, 56], let A_1 and A_2 be two real symmetric matrices (which will be respectively $A + A^T$ and $A^2 + (A^2)^T$ in our application) and denote by $P(A_1, A_2)$ the pencil constructed with A_1 and A_2 which is the set of linear combinations of A_1 and A_2 with real coefficients. $P(A_1, A_2)$ is called a d-pencil if it contains a definite matrix that is, there exist real λ and μ such that $\lambda A_1 + \mu A_2$ is positive or negative definite. Roughly speaking Finsler's theorem states the following.

THEOREM 3.1. *Let A_1 and A_2 be real symmetric matrices of order $n \geq 3$, then the following statements are equivalent:*

(i) $P(A_1, A_2)$ is a d-pencil,
(ii) $x^T A_1 x = 0 \Rightarrow x^T A_2 x > 0$.

Around the same time and independently, this problem was suggested in the U. S. by G. A. Bliss and W. T. Reid at the University of Chicago. It was solved by A. A. Albert [1] at the end of 1937 and the paper appeared in 1938. It was also considered by W. T. Reid [44]. This result was generalized by Hestenes and McShane [27] to more than two quadratic forms with applications in the calculus of variations. Let $\mathcal{Q}_i$, $i = 1, 2$, be the set $\{x \in \mathbb{R}^n \,|\, x^T A_i x = 0\}$. Dines [16] proved in 1940 that the set $\{(x^T A_1 x,\, x^T A_2 x), x \in \mathbb{R}^n\}$ is convex in the two-dimensional plane. Moreover if $\mathcal{Q}_1 \cap \mathcal{Q}_2 = \{0\}$, then this set is closed and is either the entire two-dimensional plane or an angular sector of angle less than π. The Finsler-Bliss-Albert-Reid result appeared as a corollary of one of his results.

Since then, this problem has been extensively studied (mainly for applications in optimization with quadratic constraints) and these results have been rediscovered or enhanced again and again. Among others, see the papers by Calabi [10], Hestenes [26], Donoghue [19], Uhlig [54, 55, 56], Marcus [38], Tsing and Uhlig [53], Polyak [41]. An interesting reference that is only partly devoted to this problem is Ikramov [32]. Another paper summarizing results is Hiriart-Urruty and Torki [29]. The main result is the following, as formulated in Uhlig's papers.

THEOREM 3.2. *Let A_1 and A_2 be real symmetric matrices of order $n \geq 3$ and $\mathcal{Q}_i$, for $i = 1, 2$, be the sets $\{x \in \mathbb{R}^n \,|\, x^T A_i x = 0\}$. Then the following statements are equivalent:*

(i) $P(A_1, A_2)$ is a d-pencil,
(ii) $\mathcal{Q}_1 \cap \mathcal{Q}_2 = \{0\}$,
(iii) trace$(Y A_1)$=trace$(Y A_2)$ for Y being symmetric positive semi-definite implies that $Y = 0$.

The equivalence of (i) and (ii) was formulated in this way by Calabi [10]. This is what we will mainly use for our purposes. However, condition (iii) is also directly related to our problem. Note the definition of the inner product of two real matrices A and B, $\langle A, B\rangle = \text{trace}(A^T B)$. Also remark that the values of the quadratic forms $b^T A_i b$ can be written as $\langle A_i, bb^T\rangle$. Hence, $b^T A_i b = 0$, $i = 1, \ldots, n-1$, is equivalent to the matrices A_i

ETNA
Kent State University
http://etna.math.kent.edu

being orthogonal to the positive semi-definite rank-one matrix bb^T. Therefore the existence of a stagnation vector is equivalent to the existence of a non-trivial symmetric rank-one matrix orthogonal to A_i, $i = 1, \dots, n-1$. Property (iii) implies that if there exists a stagnation vector for $n = 3$, then $P(A_1, A_2)$ is not a d-pencil.

The results of Theorem 3.2 are linked to generalizations of the field of values (or numerical range). The joint field of values of two matrices A_1 and A_2 is defined as

$$\mathcal{F}_K(A_1, A_2) = \{(\,(A_1x, x),\, (A_2x, x)\,),\, x \in K^n,\, \|x\| = 1\}, \tag{3.1}$$

where K is $\mathbb{R}$ or $\mathbb{C}$. Brickman [7] proved in 1961 that $\mathcal{F}_{\mathbb{R}}(A_1, A_2)$ is convex and also that $\mathcal{F}_{\mathbb{R}}(A_1, A_2) = \mathcal{F}_{\mathbb{C}}(A_1, A_2)$. Moreover, the two sets $\{(x^TA_1x,\, x^TA_2x), x \in \mathbb{R}^n\}$ and $\{(x^HA_1x,\, x^HA_2x), x \in \mathbb{C}^n\}$ are the same convex cone. Polyak [41] extended some of these results to three matrices.

THEOREM 3.3 (Theorem 2.1 of Polyak [41]). *Let A_1, A_2 and A_3 be real symmetric matrices of order $n \geq 3$ and $\mathcal{Q}_i$, $i = 1, 2, 3$ be the set $\{x \in \mathbb{R}^n \,|\, x^TA_ix = 0\}$. Then the following statements are equivalent:*

(i) there exist μ_1, μ_2, μ_3 such that $\mu_1A_1 + \mu_2A_2 + \mu_3A_3$ is positive definite,

(ii) $\mathcal{Q}_1 \cap \mathcal{Q}_2 \cap \mathcal{Q}_3 = \{0\}$ and the set $\{(x^TA_1x,\, x^TA_2x,\, x^TA_3x), x \in \mathbb{R}^n\}$ is an acute closed convex cone in $\mathbb{R}^3$.

However, it is interesting to note that extensions of the results of Theorem 3.2 were already considered by Dines [17, 18] in the 1940's. In [17] he looked at what is now called the real joint field of values defined by m quadratic forms and proved that it is bounded and closed. He then considered the convex extension of this set (this is what we now call the convex hull of the joint field of values). He proved that a sufficient and necessary condition to have a positive definite linear combination of the matrices A_i, $i = 1, \dots, m$, is that there is no set of vectors z_j, $j = 1, \dots, r$, such that $\sum_{j=1}^{r} m_jz_j^TA_iz_j = 0$, $i = 1, \dots, m$, with positive coefficients m_j. He also extended these results to the positive semi-definite case. In [18] Dines extended the equivalence of (i) and (iii) in Theorem 3.2 to m quadratic forms. He proved that having a definite linear combination of the matrices A_i, $i = 1, \dots, m$, is equivalent to having every matrix B orthogonal to the matrices A_i, $\langle A_i, B\rangle = 0$, $i = 1, \dots, m$, indefinite. There exists a definite symmetric matrix B orthogonal to A_i, $i = 1, \dots, m$, if and only if every linear combination of the A_is is indefinite. This type of results was also extended to semi-definite matrices. However, the matrix B is not necessarily rank-one.

There does not seem to exist direct extensions of the result of Theorem 3.3 to more than three matrices. This is probably because the set $\{(x^TA_1x, \dots, x^TA_mx)),\;\; x \in \mathbb{R}^n\}$ is not always a closed convex cone for $m > 3$. However, there exist a few generalizations of Finsler's theorem; see Hamburger [25], Arutyunov [4], Ai, Huang, and Zhang [2]. For the joint field of values with m matrices and $x \in \mathbb{C}^n$, see Fan and Tits [20], Gutkin, Jonckheere, and Karow [24, Proposition 2.10], and Chien and Nakazato [12].

3.2. The case $n = 2$. Let us consider real 2×2 matrices. This case can be solved easily. We have only one orthogonality condition $b^TAb = 0$, to which we add $b^Tb = 1$. Such a vector b is called an isotropic vector in the literature; see [11, 13, 40, 57] for algorithms to compute isotropic vectors. However, the problem is simpler when $n = 2$. As we have defined before, let A_1 be twice the symmetric part of A. Then $A_1 = Q\Lambda Q^T$, $y = Q^Tb$, with Q orthogonal and Λ diagonal and

$$b^TAb = 0 \Leftrightarrow b^TA_1b = 0 \Leftrightarrow b^TQ\Lambda Q^Tb = 0 \Leftrightarrow y^T\Lambda y = 0.$$

To have non-trivial solutions, the matrix A_1 has to be indefinite. So there must be one positive eigenvalue λ_1 and one negative eigenvalue $-\lambda_2$, and the condition $b^TA_1b = 0$ reads

$$\lambda_1y_1^2 - \lambda_2y_2^2 = 0 \Leftrightarrow (\sqrt{\lambda_1}y_1 - \sqrt{\lambda_2}y_2)(\sqrt{\lambda_1}y_1 + \sqrt{\lambda_2}y_2) = 0.$$

The solution set of this equation is the union of two lines passing through the origin with respective slopes $\pm\sqrt{\lambda_1}/\sqrt{\lambda_2}$. The eigenvalues of A_1 are

$$\lambda = a_{1,1} + a_{2,2} \pm [(a_{1,1} - a_{2,2})^2 + a_{121}^2]^{1/2},$$

where $a_{121} = a_{1,2} + a_{2,1}$. To obtain the solutions having $y^T y = 1$, we have to intersect the lines with the unit circle. This gives four solutions with

$$y_1 = \pm\frac{\sqrt{\lambda_2}}{\sqrt{\lambda_1 + \lambda_2}}, \quad y_2 = \pm\frac{\sqrt{\lambda_1}}{\sqrt{\lambda_1 + \lambda_2}}.$$

Then we have to rotate the solutions to obtain $b = Qy$. The solutions are entirely determined by the eigenvalues and eigenvectors of A_1. The only condition to have solutions for $n = 2$ is to have A_1 indefinite. This happens if $a_{121}^2 - 4a_{1,1}a_{2,2} > 0$.

3.3. The case $n = 3$. For real matrices of order 3 we have a polynomial system of degree 2 with 3 equations,

$$b^T(A + A^T)b = b^T A_1 b = 0,\; b^T(A^2 + (A^2)^T)b = b^T A_2 b = 0,\; b^T b = 1.$$

There is a real stagnation vector if and only if $(0, 0)$ is in the (real) joint field of values defined in (3.1) with $K = \mathbb{R}$. The set $\mathcal{F} \equiv \mathcal{F}_{\mathbb{R}}$ coincides with the classical numerical range $W(B)$ of $B = A_1 + \imath A_2$; see [30]. Since A_1 and A_2 are real and symmetric, the matrix B is symmetric but not Hermitian. Hence, the set $\mathcal{F}$ is not symmetric with respect to the real axis. Many results are known about the numerical range that can be used to study the properties of $\mathcal{F}$; see, for instance, [30, 35, 43]. In particular, $\mathcal{F}$ is a compact convex set in the two-dimensional plane. Therefore, it is closed and bounded. The result on the convexity of $W(B)$ is the celebrated Toeplitz–Hausdorff theorem that was proved in 1918. Since A_1 and A_2 are the Hermitian and skew-Hermitian part of B and symmetric, we have that

$$\begin{aligned} Re\,[W(B)] &= W(A_1) = [\lambda_{min}(A_1), \lambda_{max}(A_1)], \\ Im\,[W(B)] &= W(A_2) = [\lambda_{min}(A_2), \lambda_{max}(A_2)]. \end{aligned}$$

Hence, $\mathcal{F}$ is enclosed in the rectangle $[\lambda_{min}(A_1), \lambda_{max}(A_1)] \times [\lambda_{min}(A_2), \lambda_{max}(A_2)]$. The boundary of the numerical range $W(B) \equiv \mathcal{F}$ can be sketched by considering the matrices $B_\theta = e^{\imath\theta}B = (\cos(\theta) + \imath\sin(\theta))B$ for $\theta \in [0, 2\pi]$; see [34]. The Hermitian part of B_θ is $\cos(\theta)A_1 - \sin(\theta)A_2$. Note that this matrix is symmetric. Let λ_{max}^θ be its largest eigenvalue and x_θ be the associated eigenvector. Then an (inner) polygonal approximation of $W(B)$ is obtained by linking the points $x_\theta^T B x_\theta$ for $0 = \theta_1 < \theta_2 < \cdots < \theta_p = 2\pi$. Since x_θ is real, the coordinates of these points are $x_\theta^T A_1 x_\theta$ and $x_\theta^T A_2 x_\theta$. Note that we could have as well considered the matrix $\cos(\theta)A_1 + \sin(\theta)A_2$ and also the smallest eigenvalue instead of the largest.

The boundary of $\mathcal{F}$ was also characterized in Polyak [41] without reference to $W(B)$. We will use the following definition for the points on the boundary of $\mathcal{F}$,

$$\{(x_\theta^T A_1\, x_\theta,\; x_\theta^T A_2\, x_\theta),\; \theta \in [0, 2\pi]\}. \tag{3.2}$$

We will see in the numerical experiments that in some examples, there is one point on the boundary of $\mathcal{F}$ in the vicinity of which $\mathcal{F}$ looks like a sector. Those points are called "corners" or "conical points". It is well-known that the coordinates of those points are the real and imaginary parts of an eigenvalue of B; see [30, 43]. Moreover, the geometric multiplicity of that eigenvalue is equal to its algebraic multiplicity.

ETNA
Kent State University
http://etna.math.kent.edu

In many 3×3 examples, the boundary of $\mathcal{F}$ has almost "flat" portions. This phenomenon has been studied in [8, 45]. Conditions are given in [35] for the field of values of a 3×3 matrix to be the convex hull of a point and an ellipse. Roughly speaking, when the matrix is not normal, the field of values can have this shape or it can be an ellipse or an ovular set.

In general, there is no known necessary and sufficient condition for the origin to be a point of a numerical range. However, for our particular problem, we have the following necessary and sufficient condition for the non-existence of a stagnation vector.

THEOREM 3.4. *Let A be a real matrix of order $n = 3$. There is no real stagnation vector if and only if there exist real numbers λ and μ such that $\lambda(A + A^T) + \mu(A^2 + (A^2)^T)$ is definite.*

Proof. The result is a direct consequence of Theorem 3.2. □

Of course now, the question is to know when is $\lambda A_1 + \mu A_2$ definite. For our small problem of order 3, this can be done by direct examination as we will see in the numerical experiments. But this question has been investigated for matrices of order n by several researchers. Uhlig published several papers in the 1970's; see [54, 55, 56]. An algorithm was proposed by Crawford and Moon [14, 15] to compute such a pair (λ, μ); see also [32]. This problem has also been considered by Higham, Tisseur, and Van Dooren [28, Algorithm 2.4]. The following result is well-known. It is in the same spirit as results of O. Taussky [51].

LEMMA 3.5. *Let A_1 and A_2 be real symmetric matrices such that there exist λ and μ with $\lambda A_1 + \mu A_2$ (say) positive definite. Then there exists a real nonsingular matrix X such that $\Omega = X^T A_1 X$ and $\Gamma = X^T A_2 X$ are diagonal.*

The ratios of the diagonal elements of $X^T A_1 X$ and $X^T A_2 X$ are the eigenvalues of the pencil (A_1, A_2). Note that in order to compute X, we have to know a pair (λ, μ) such that $\lambda A_1 + \mu A_2$ is positive definite. Details on the region where $\lambda A_1 + \mu A_2$ is definite were given by Uhlig [56].

THEOREM 3.6 (Theorem 1.1 of Uhlig [56]). *Let (A_1, A_2) be a d-pencil. Let X be such that $X^T A_1 X = diag(\gamma_i)$ and $X^T A_2 X = diag(\omega_i)$.*

If there exist indices i, j such that $\gamma_i \gamma_j < 0$, then

1.
$$\max_{\gamma_i > 0} \frac{\omega_i}{\gamma_i} < \max_{\gamma_i < 0} \frac{\omega_i}{\gamma_i},$$
and $\omega_i < 0$ whenever $\gamma_i = 0$, or

2.
$$\min_{\gamma_i > 0} \frac{\omega_i}{\gamma_i} > \min_{\gamma_i < 0} \frac{\omega_i}{\gamma_i},$$
and $\omega_i > 0$ whenever $\gamma_i = 0$.

In case that all the γ_i's have the same sign

3. *either $\omega_i < 0$ whenever $\gamma_i = 0$ or $\omega_i > 0$ whenever $\gamma_i = 0$.*

If we know γ_i and ω_i, we can compute the boundary of the region where $\lambda A_1 + \mu A_2$ is (say) positive definite.

THEOREM 3.7 (Theorem 1.2 of Uhlig [56]). *Let (A_1, A_2) be a d-pencil. Using the notation of Theorem 3.6, the matrix $\lambda A_1 + \mu A_2$ is positive definite if and only if (the cases correspond to Theorem 3.6)*

1.
$$-\left(\max_{\gamma_i > 0} \frac{\omega_i}{\gamma_i}\right)^{-1} < \frac{\lambda}{\mu} < -\left(\max_{\gamma_i < 0} \frac{\omega_i}{\gamma_i}\right)^{-1},$$

ETNA
Kent State University
http://etna.math.kent.edu

2.

$$-\left(\min_{\gamma_i>0}\frac{\omega_i}{\gamma_i}\right)^{-1} > \frac{\lambda}{\mu} > -\left(\min_{\gamma_i<0}\frac{\omega_i}{\gamma_i}\right)^{-1},$$

3. If all $\gamma_i \geq 0$, then

$$-\left(\max_{\gamma_i>0}\frac{\omega_i}{\gamma_i}\right)^{-1} < \frac{\lambda}{\mu} < 0 \text{ if } \omega_i < 0 \text{ whenever } \gamma_i = 0,$$

and

$$0 < \frac{\lambda}{\mu} < -\left(\min_{\gamma_i>0}\frac{\omega_i}{\gamma_i}\right)^{-1} \quad \text{if } \omega_i > 0 \text{ whenever } \gamma_i = 0,$$

and if all $\gamma_i \leq 0$, then

$$-\left(\max_{\gamma_i>0}\frac{\omega_i}{\gamma_i}\right)^{-1} > \frac{\lambda}{\mu} > 0 \text{ if } \omega_i < 0 \text{ whenever } \gamma_i = 0,$$

and

$$0 > \frac{\lambda}{\mu} > -\left(\min_{\gamma_i>0}\frac{\omega_i}{\gamma_i}\right)^{-1} \quad \text{if } \omega_i > 0 \text{ whenever } \gamma_i = 0.$$

Note that these inequalities define intersections of half-planes in the $(\lambda,\,\mu)$ plane. The following two algorithms simplify the problem by using only one parameter (like it is done for studying the boundary of $\mathcal{F}$) instead of λ and μ. A sketch of the algorithm proposed by Crawford and Moon [15] to produce a pair (λ, μ) such that $\lambda A_1 + \mu A_2$ is positive definite is the following; see also [32]. They consider the function

$$g(x) = \frac{(A_1x,x) + \imath(A_2x,x)}{|(A_1x,x) + \imath(A_2x,x)|},\ x \in \mathbb{C}^n.$$

The function g is not defined if $(A_1x,x) = 0$ and $(A_2x,x) = 0$ simultaneously. Otherwise, the values of g belong to an arc of the unit circle. The algorithm is a bisection method to locate the end points of this arc. Having an approximation $[a,\,b]$ of the arc, one considers the mid-point c and the corresponding angle t formed by the imaginary axis and the segment $[0,\,c]$. If $B(t) = \sin(t)A_1 + \cos(t)A_2$ is positive definite, a pair $\lambda = \sin(t), \mu = \cos(t)$ has been found. Otherwise a vector x such that $(B(t)x,x) \leq 0$ is computed as well as $d = g(x)$. If d belongs to the arc $(-a,\,b]$, then $b = d$; if d belongs to the arc $[a,\,-b)$, then $a = d$. If none of these conditions is satisfied, the algorithm has failed. The vector x is computed by using the partial Cholesky decomposition of $B(t)$ using the upper triangular matrix R_k of the maximal positive definite principal matrix of $B(t)$.

The algorithm of Higham, Tisseur, and Van Dooren [28, Algorithm 2.4, p. 462] computes the Crawford number

$$\gamma(A_1,A_2) = \min\sqrt{(A_1x,x)^2 + (A_2x,x)^2}\ \text{ with } x \in \mathbb{C}^n,\ \|x\| = 1.$$

It also gives the answer to our problem. Let $B = A_1 + \imath A_2$ and compute the eigenvalues of the quadratic polynomial $P(\lambda) = B - \lambda^2 B^H$, where B^H is the conjugate transpose. If there are $2n$ eigenvalues of modulus one, for each eigenvalue $\lambda_j = e^{\imath\theta_j}$ we compute an eigenvector v of $A_\theta = \cos(\theta_j)A_1 + \sin(\theta)A_2$ and the sign of $(B_\theta v, v)$, where $B_\theta = \cos(\theta_j)A_2 - \sin(\theta)A_1$,

ETNA
Kent State University
http://etna.math.kent.edu

which gives the sign of the derivative. If there are n consecutive points θ_j with the same sign, there exist λ and μ such that $\lambda A_1 + \mu A_2$ is definite. They are given by the angles in the interval starting at the nth point with the same sign and ending at the next θ_j. Note that there exist more general algorithms that will be described in the next subsection.

Finally, we can obtain conditions for the existence of stagnation vectors using the eigenvalues of the matrix pencil $(A_1,\ A_2)$. Let

$$A_1 y^i = \lambda_i A_2 y^i, \tag{3.3}$$

be the generalized eigenvalues and eigenvectors that are those of $A_1^{-1}A_2$ since A_1 is nonsingular.

LEMMA 3.8. *Let A_1 and A_2 be real symmetric matrices. Let λ_i be the eigenvalues such that $A_1 x^i = \lambda_i A_2 x^i$ and $\hat{\lambda}_i$ the eigenvalues such that $A_1 y^i = \hat{\lambda}_i(\lambda A_1 + \mu A_2)y^i$. Then for $\lambda_i \neq 0$, the eigenvectors of both pencils are the same and the eigenvalues are related by*

$$\lambda_i(1 - \lambda\hat{\lambda}_i) = \mu\hat{\lambda}_i. \tag{3.4}$$

Proof. Assume $\lambda_i \neq 0$ and $A_1 x^i = \lambda_i A_2 x^i$. Then x^i is an eigenvector of the pencil $(A_1, \lambda A_1 + \mu A_2)$ and we have

$$A_1 x^i = \hat{\lambda}_i(\lambda A_1 + \mu A_2)x^i = \hat{\lambda}_i(\lambda\lambda_i + \mu)A_2 x^i.$$

It gives $\hat{\lambda}_i(\lambda\lambda_i + \mu) = \lambda_i$. □

THEOREM 3.9. *Let A_1 and A_2 be real symmetric matrices of order $n \geq 3$. If there are complex eigenvalues λ_i in (3.3), then there is no real λ and μ such that $\lambda A_1 + \mu A_2$ is definite. Of course, if there exist real λ and μ such that $\lambda A_1 + \mu A_2$ is definite, then the eigenvalues in (3.3) are real.*

Proof. These results were stated in Polyak [41] without proof. The proof is by contradiction. Let us assume that there exist real λ and μ such that $\lambda A_1 + \mu A_2$ is positive definite. First, we reduce the problem to a pencil $(A_1, \hat{A}_2)$ with $\hat{A}_2 = \lambda A_1 + \mu A_2$ being symmetric positive definite. Using Lemma 3.8, the eigenvalues are related by (3.4). Now, since in the pencil $(A_1, \hat{A}_2)$ the matrix $\hat{A}_2$ is positive definite, the eigenvalues are all real. This can be seen using the Cholesky factorization of $\hat{A}_2 = LL^T$ with L being lower triangular. We have

$$A_1 y^i = \hat{\lambda}_i \hat{A}_2 y^i = \hat{\lambda}_i LL^T y^i \Rightarrow (L^{-1}A_1L^{-T})(L^T y^i) = \hat{\lambda}_i(L^T y^i).$$

Since the matrix $(L^{-1}A_1L^{-T})$ is symmetric, the eigenvalues $\hat{\lambda}_i$ are real. This implies that all the eigenvalues λ_i of the original pencil are also real, which is contradictory to our hypothesis. □

The converse of Theorem 3.9 is not true. There exist pencils having all the eigenvalues real for which there is no λ and μ such that $\lambda A_1 + \mu A_2$ is definite. However, we have the following sufficient condition for the existence of stagnation vectors.

THEOREM 3.10. *Let A be a real matrix of order $n = 3$. If complex eigenvalues of the pencil $(A + A^T, A^2 + (A^2)^T)$ exist, then A has real stagnation vectors.*

3.4. The case $n = 4$. Theorem 3.3 solves the stagnation problem for $n = 4$. In this case we have the polynomial system

$$b^T(A + A^T)b = b^TA_1b = 0, \qquad b^T(A^2 + (A^2)^T)b = b^TA_2b = 0,$$
$$b^T(A^3 + (A^3)^T)b = b^TA_3b = 0, \qquad b^Tb = 1.$$

ETNA
Kent State University
http://etna.math.kent.edu

THEOREM 3.11. *Let A be a real matrix of order $n = 4$. Then there is no real stagnation vector and the set $F = \{(x^TA_1x,\ x^TA_2x,\ x^TA_3x),\ x \in \mathbb{R}^n\}$ is convex if and only if there exist real μ_i, $i = 1, 2, 3$, such that $\mu_1A_1 + \mu_2A_2 + \mu_3A_3$ is definite.*

Proof. The set F is obviously a cone with vertex 0. It has been proved to be closed by Dines [17] and we assume that it is convex. In Theorem 3.3, F is assumed to be an acute cone, but this was because equivalence was sought with having a linear combination that is positive definite. Considering the proof of Theorem 3.3, one can handle the cases for positive and negative definite matrices separately. ∎

An interesting question (which seems to be an open one) is to know if the hypothesis on the convexity of F is necessary for our problem. If one looks at the proof of Theorem 3.3, which relies on [36], convexity is not absolutely needed. It is just necessary to have the existence of a linear functional which is (say) strictly positive on F. The set F has to be on one side of a plane containing the origin. It is known that $\mathcal{F}_{\mathbb{C}}$ is convex. This should imply the convexity of

$$F_{\mathbb{C}} = \{(x^HA_1x,\ x^HA_2x,\ x^HA_3x),\ x \in \mathbb{C}^n\}.$$

We also have $F \subseteq F_{\mathbb{C}}$. Therefore, F is contained in a convex cone in $\mathbb{R}^3$. This should be enough to prove that there exists a linear functional which is strictly positive on F (e.g. F is on one side of a plane passing through the origin).

To our knowledge, there is no characterization (like the one of Uhlig for two matrices) of the regions in μ_1, μ_2, μ_3 where $\mu_1(A + A^T) + \mu_2(A^2 + (A^2)^T) + \mu_3(A^3 + (A^3)^T)$ is positive or negative definite in the literature. However, the boundary of the joint field of values with the constraint $\|x\| = 1$ is known; see [37]. One considers two angles $0 \le \theta, \phi \le 2\pi$ and the matrices $T_\theta = \cos(\theta)A_1 + \sin(\theta)A_2$ and $B_{\theta,\phi} = \sin(\phi)T_\theta + \cos(\phi)A_3$. Let $x_{\theta,\phi}$ be the eigenvector corresponding to the largest (or smallest) eigenvalue of $B_{\theta,\phi}$. The coordinates of the boundary points are given by $(x_{\theta,\phi}^TA_1x_{\theta,\phi},\ x_{\theta,\phi}^TA_2x_{\theta,\phi},\ x_{\theta,\phi}^TA_3x_{\theta,\phi})$. Psarrakos [42] used this characterization to propose an algorithm to determine if a triple of Hermitian matrices is definite. In case the origin is in the joint field of values, he also computed the next definite triple with a given Crawford number.

There exist more general algorithms to compute parameters $\mu_1, \dots, \mu_m$ (when they exist) such that $A(\mu) = \sum_{j=1}^m \mu_jA_j$ is (say) positive definite. An algorithm denoted as PC was proposed by Tong, Iujiro, and Liu [52]. Starting from a vector $\mu = (\mu_j)$, $j = 1, \dots, m$, this method iteratively updates μ by $\mathrm{d}\mu$ such that

$$\mathrm{d}\mu = \frac{(x^TA_1x, \dots, x^TA_mx^T)^T}{\|(x^TA_1x, \dots, x^TA_mx^T)^T\|},$$

where x is the eigenvector corresponding to the smallest eigenvalue of $B = \sum_{j=1}^m \mu_jA_j$. The iterations are stopped when a positive definite matrix B has been found or when the maximum number of iterations has been reached. If there is a set of coefficients such that $A(\mu)$ is positive definite, the algorithm converges but sometimes very slowly. Zaidi [58] proposed the Positive Definite Combination (PDC) algorithm. His goal was to obtain a vector of coefficients μ such that the smallest eigenvalue of $A(\mu)$ is larger than a given $\sigma \ge 0$. Finding such a matrix $A(\mu)$ is formulated as an optimization problem with constraints,

$$\min_{\mu, C} \|B - C\|_F,$$

subject to $B = \sum_{j=1}^m \mu_jA_j$ and $C \in S_\sigma^+$ where

$$S_\sigma^+ = \{QDQ^T \,|\, D = \sigma I + \Delta; Q \text{ orthogonal}\},$$

ETNA
Kent State University
http://etna.math.kent.edu

and Δ being a diagonal matrix with diagonal entries $\delta_j \geq 0$. The minimization problem is solved by an alternating algorithm, minimizing the function successively with respect to μ and to C. Using an arbitrary vector of coefficients μ, we start from $B = \sum_{j=1}^{m} \mu_j A_j = U\Lambda U^T$ with U orthogonal and $\Lambda = (\lambda_j)$ diagonal. We choose an arbitrary diagonal matrix Δ with a positive diagonal and set $C = U(\sigma I + \Delta)U^T$. Then, as long as the smallest eigenvalue λ_{min} is negative, we iterate: we set

$$D = diag(\max(\lambda_i, \sigma)),$$

and $B = P(UDU^T)$ and we compute a new eigendecomposition $B = U\Lambda U^T$. Finally the new matrix C is $C = UDU^T$. The matrix $P(M)$ is the orthogonal projection of M on the subspace spanned by the matrices A_j, $j = 1, \ldots, m$. This can be computed by vectorization. Let $vec(M)$ be the vector of length n^2 defined by stacking the columns of M and $\mathcal{A} = \begin{bmatrix} vec(A_1) & \cdots & vec(A_m) \end{bmatrix}$. Conversely, let $mat(v)$ be the matrix of order n constructed from the vector v of length n^2. Then

$$P(M) = mat(\mathcal{A}(\mathcal{A}^T\mathcal{A})^{-1}\mathcal{A}^T vec(M)).$$

Note that there are some misprints in [58]. A similar algorithm was introduced by Cai, Peng, and Zhang [9]. However, they added a constraint on the vector μ, for instance $\|\mu\| = 1$. This makes the projection phase more difficult, but it allows to determine if there is no positive definite linear combination when $\sigma > 0$. Recently, some other algorithms have been proposed by Huhtanen and Seiskari [31].

4. Numerical examples. In this section we describe some numerical experiments for real matrices with small values of n to illustrate the theoretical results of the previous sections. When n is small, real stagnation vectors b can be computed by solving the polynomial system (2.1). This can be done in several ways since many methods are available in the literature although mainly for polynomial systems with integer or rational coefficients. However, we are interested in polynomial systems with real coefficients. One possibility is to use homotopy; see Allgower and Georg [3] and Sommese and Wampler [49], particularly the Matlab software Homlab available on the Web. Another possibility is a method proposed by Auzinger and Stetter; see [5, 6] and also Stetter [50]. Finally, we devised an elimination method related to Gröbner bases, specially tailored to the system (2.1). Unfortunately, all these methods do not allow to solve large systems due to too large computing times for some of them and to numerical instabilities for others. The largest systems we were able to solve reliably with IEEE double precision were for $n = 6$.

4.1. The case $n = 2$. Let us consider a 2×2 (rounded) real random matrix,

$$A = \begin{bmatrix} -0.432565 & 0.125332 \\ -1.66558 & 0.287676 \end{bmatrix}.$$

This matrix has two complex conjugate eigenvalues and 0 is in the field of values of A. This is the first matrix we get using `randn(2,2)` when starting Matlab 7. The matrix $A_1 = A + A^T$ is indefinite since its eigenvalues are $\lambda_1 = 1.55544$ and $-\lambda_2$ with $\lambda_2 = 1.84522$.

The 4 stagnation vectors of unit norm (rounded) are

$$\begin{pmatrix} 0.175136 \\ 0.984544 \end{pmatrix}, \begin{pmatrix} 0.96604 \\ -0.258394 \end{pmatrix}, \begin{pmatrix} -0.96604 \\ 0.258394 \end{pmatrix}, \begin{pmatrix} -0.175136 \\ -0.984544 \end{pmatrix}.$$

The values of $b^T A b$ for the 4 solutions are

$$-5.55112\ 10^{-17}, \quad 3.79904\ 10^{-16}, \quad 3.24393\ 10^{-16}, \quad 5.55112\ 10^{-17}.$$

The study of the problem using the eigenvalues and eigenvectors of A_1 is illustrated in Figure 4.1. The blue lines are the solutions of $\lambda_1 y_1^2 - \lambda_2 y_2^2 = 0$. Then they are rotated using the matrix of the eigenvectors of A_1 to obtain the red lines. The green intersections with the circle of center $(0, 0)$ and radius one give the solutions (green circles).

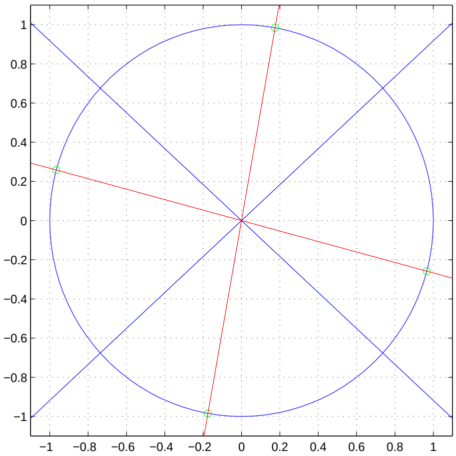

FIG. 4.1. *Example with* $n = 2$.

4.2. The case $n = 3$. This is a very interesting case because it is not as trivial as $n = 2$, and we can still visualize what happens. The maximum number of real solutions is 8. Of course, the most obvious way to know if there are real stagnation vectors is to solve the polynomial system (2.1). If all the solutions are complex, there is no real stagnation vector. However, we will numerically illustrate the theoretical results of the previous sections. Let us consider the following example with a random matrix

$$A = \begin{bmatrix} 0.614463 & 0.591283 & -1.00912 \\ 0.507741 & -0.643595 & -0.0195107 \\ 1.69243 & 0.380337 & -0.0482208 \end{bmatrix}. \tag{4.1}$$

This matrix has a pair of complex conjugate eigenvalues and a real one. The matrices $A_1 = A + A^T$ and $A_2 = A^2 + (A^2)^T$ are indefinite. The matrix $A_1^{-1} A_2$ has complex eigenvalues and therefore according to Theorem 3.10 there is no λ, μ such that $\lambda A_1 + \mu A_2$ is definite. Hence, there exist real stagnations vectors. The system (2.1) has 8 solutions but only 4 of them are real. The 4 real solutions (rounded) are

$$\begin{pmatrix} -0.47173 \\ -0.867275 \\ 0.159076 \end{pmatrix}, \begin{pmatrix} 0.47173 \\ 0.867275 \\ -0.159076 \end{pmatrix}, \begin{pmatrix} -0.198789 \\ -0.81042 \\ -0.551092 \end{pmatrix}, \begin{pmatrix} 0.198789 \\ 0.81042 \\ 0.551092 \end{pmatrix}.$$

The values of the quadratic forms for the solutions are of the order of 10^{-15}. For $n = 3$ we can visualize the quadrics defined by equation (2.1). The system for real solutions is equivalent to

$$b^T A_1 b = 0, \quad b^T A_2 b = 0, \quad b^T b = 1.$$

The matrix A_1 has two negative and one positive eigenvalues, but we can change the signs to have only one negative eigenvalue. If we diagonalize A_1 we have the equation

$$\lambda_1 x^2 + \lambda_2 y^2 - \lambda_3 z^2 = 0,$$

ETNA
Kent State University
http://etna.math.kent.edu

with $\lambda_i > 0$, $i = 1, 2, 3$, and λ_3 corresponding to the negative eigenvalue. This is the equation of a cone with an elliptical section whose axis is the z-axis. We are interested in the intersection of this cone with the unit sphere. Eliminating z in the previous equation, the intersection is defined by

$$(\lambda_1 + \lambda_3)x^2 + (\lambda_2 + \lambda_3)y^2 = \lambda_3, \quad z^2 = 1 - x^2 - y^2.$$

The first equation defines an ellipse of semi-axes $\sqrt{\lambda_3/(\lambda_1 + \lambda_3)}$ and $\sqrt{\lambda_3/(\lambda_2 + \lambda_3)}$. Then this ellipse is "projected" onto the unit sphere. The intersection is the union of two smooth closed curves on the surface of the sphere. They are symmetric with respect to the origin. This is shown in Figure 4.2. We only show the upper half of the cone. The intersection with the unit sphere is the blue thick curve. Then we use the eigenvectors of A_1 to rotate the cone and therefore also the blue curves. Note that angles and distances are preserved in this rotation. The result is displayed in Figure 4.3.

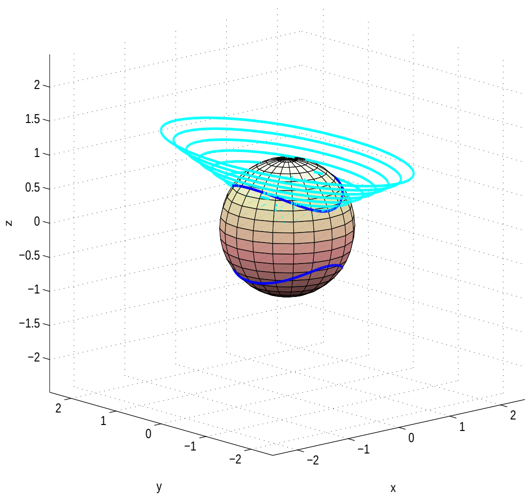

FIG. 4.2. $n = 3$, *quadratic form for* H *in* x, y, z *frame, matrix (4.1).*

Then we consider the cone for A_2. Note that the values $\sqrt{\lambda_3/(\lambda_1 + \lambda_3)}$ and $\sqrt{\lambda_3/(\lambda_2 + \lambda_3)}$ are related to the aperture of the cone. It these numbers are small, the aperture of the cone is small. This happens if $\lambda_1 + \lambda_3$ and $\lambda_2 + \lambda_3$ are large compared to λ_3 (the absolute value of the negative eigenvalue). The fact that the two cones intersect or not depend on their apertures and also on their respective positions. Nevertheless, generally if at least one of the two cones has a small aperture, it is likely that there is no intersection (although this possibility is not ruled out). It turns out that for this example, after the rotations, the two cones intersect. Figure 4.4 shows the intersections (green circles) of the blue (for A_1) and red (for A_2) curves on the surface of the unit sphere. We see two intersections. One of the red curves intersect only one of the blue curves. The two other intersections are located on the other side of the sphere. There are obtained by symmetry. The unfortunately missing parts of the curves are due to an artifact during the translation from Matlab to Postscript.

To obtain an insight about the number of solutions, we may look at the two cones in the frame defined by the eigenvectors of A_1. If the two cones intersect, the ellipses defined by taking z constant must intersect. Twice the number of intersections gives us the number of

ETNA
Kent State University
http://etna.math.kent.edu

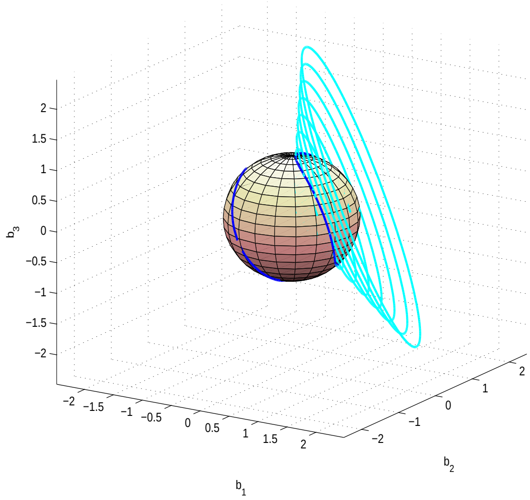

FIG. 4.3. $n = 3$, *quadratic form for* H *in* b_1, b_2, b_3 *frame, matrix* (4.1).

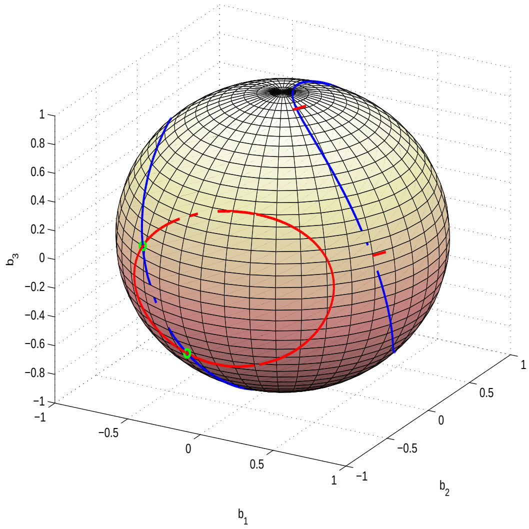

FIG. 4.4. $n = 3$, *the solutions (green circles) in* b_1, b_2, b_3 *frame, matrix* (4.1).

real stagnation vectors. Let us take $z = 1$. The equation for A_1 is

$$\lambda_1 x^2 + \lambda_2 y^2 = \lambda_3. \tag{4.2}$$

Let $K = A - A^T$ and Q be the matrix of eigenvectors of A_1. It is easy to see that $A_2 = (A_1^2 + K^2)/2$. Here K^2 is a singular matrix with two negative eigenvalues. The equation for A_2 is

$$\lambda_1^2 x^2 + \lambda_2^2 y^2 + \lambda_3^2 + \begin{pmatrix} x & y & 1 \end{pmatrix} Q^T K^2 Q \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = 0. \tag{4.3}$$

ETNA
Kent State University
http://etna.math.kent.edu

Now, the problem is to know if the two ellipses, defined by (4.2) and (4.3), intersect and, eventually, how many intersection points we have. The equations of the ellipses can be considered as quadratic polynomials in x with coefficients that are polynomials in y; see, for instance, [23]. These two quadratic polynomials in x have a common root if and only if the discriminant is zero. Let $C = Q^T K^2 Q$,

$$\begin{aligned}
&\alpha_0 = 2\lambda_1 c_{1,2}, && \alpha_1 = \lambda_1(\lambda_2^2 + c_{2,2}) - \lambda_2(\lambda_1^2 + c_{1,1}), && \alpha_2 = 2\lambda_1 c_{1,3},\\
&\alpha_3 = 2\lambda_1 c_{2,3}, && \alpha_4 = \lambda_1(\lambda_3^2 + c_{3,3}) + \lambda_3(\lambda_1^2 + c_{1,1}), && \alpha_5 = -2\lambda_2 c_{1,2},\\
&\alpha_6 = 0, && \alpha_7 = -2\lambda_3 c_{1,2}, && \alpha_8 = 2\lambda_1 c_{1,3},\\
&\alpha_9 = 0, && \alpha_{10} = 2\lambda_3 c_{1,3},
\end{aligned}$$

and

$$\begin{aligned}
\beta_0 &= \alpha_2\alpha_{10} - \alpha_4^2,\\
\beta_1 &= \alpha_0\alpha_{10} + \alpha_2(\alpha_7 + \alpha_9) - 2\alpha_3\alpha_4,\\
\beta_2 &= \alpha_0(\alpha_7 + \alpha_9) + \alpha_2(\alpha_6 - \alpha_8) - \alpha_3^2 - 2\alpha_1\alpha_4,\\
\beta_3 &= \alpha_0(\alpha_6 - \alpha_8) + \alpha_2\alpha_5 - 2\alpha_1\alpha_3,\\
\beta_4 &= \alpha_0\alpha_5 - \alpha_1^2.
\end{aligned}$$

Then we have

$$\beta_4 y^4 + \beta_3 y^3 + \beta_2 y^2 + \beta_1 y + \beta_0 = 0. \tag{4.4}$$

The roots of this quartic polynomial give the y coordinates of the intersections. Then the x coordinates can be computed by (4.2). Unfortunately, it does not seem possible to obtain analytic expressions for the roots. The number of real solutions, which can be obtained by Sturm's theorem, is half the number of the real stagnation vectors.

Of course, for this example the origin is in the fields of values of A and A^2. Otherwise we would not have a solution. However, it is more interesting to look at the joint field of values $\mathcal{F}_{\mathbb{R}}(A_1, A_2)$ in the two-dimensional plane. We can use brute force to visualize this set by plotting values for random unit vectors x as it is shown in Figure 4.5 with blue plus signs; here we have 400 points. Note that, even though the points are not uniformly distributed, this gives a good idea of the shape of the joint field of values. The green box is given by the eigenvalues of A_1 and A_2 whose pairs are displayed as light blue stars. The red star is $(0, 0)$ which is within the joint field of values, meaning that there exist real stagnation vectors. The green curve and stars show the boundary of the joint field of values as given by (3.2). They were obtained with a uniform mesh in $[0,\ 2\pi]$. In many examples with random matrices of order 3, the joint field of values has such a "triangular" shape. The shape is more or less the convex hull of an ellipse and a point outside. The "corners" are eigenvalues of $A_1 + \imath A_2$ and also close to pairs of eigenvalues of A_1 and A_2. Some blue crosses seem outside the boundary but this is because we do not have enough points on some parts of the boundary. We remark that the computed boundary points are concentrated around the "corners" of the two-dimensional set. There are also two portions of the boundary that look like straight lines without any discretization points. Remember that the green points are given by pairs $(x(t)^T A_1 x(t), x(t)^T A_2 x(t))$ where $x(t)$ is the eigenvector corresponding to the smallest eigenvalue of $A_t = \cos(t)A_1 + \sin(t)A_2$. Figure 4.6 shows the three eigenvalues of A_t as functions of t in $[0,\ 2\pi]$. The smallest (resp. largest) eigenvalue is always negative (resp. positive). We see that A_t is never definite. Figure 4.7 displays the values $x(t)^T A_1 x(t)$ (green curve) and $x(t)^T A_2 x(t))$ (magenta curve) as functions of t. There are values of t for which there is a large increase (or decrease) in the functions. This corresponds to the parts of the boundary of the joint field of values that look like straight lines. Note that these values of t are some of the ones for which two eigenvalues of A_t are close to each other. The parts of the boundary where we have an accumulation of points correspond to the "flat" parts of Figure 4.7.

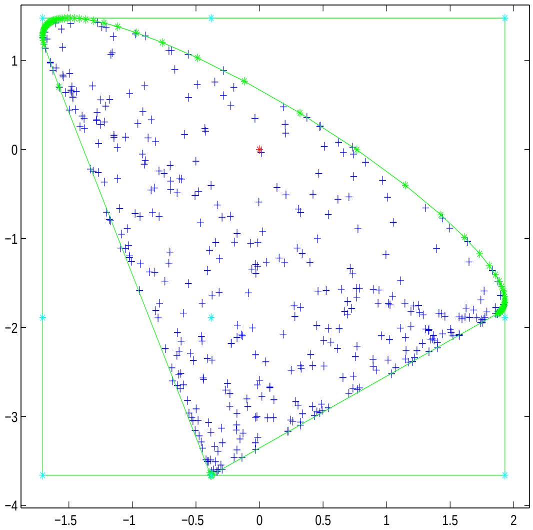

FIG. 4.5. $n = 3$, *joint field of values*, $\{ (x^TA_1x, x^TA_2x),\ x \in \mathbb{R}^3,\ \|x\| = 1 \}$, *matrix (4.1).*

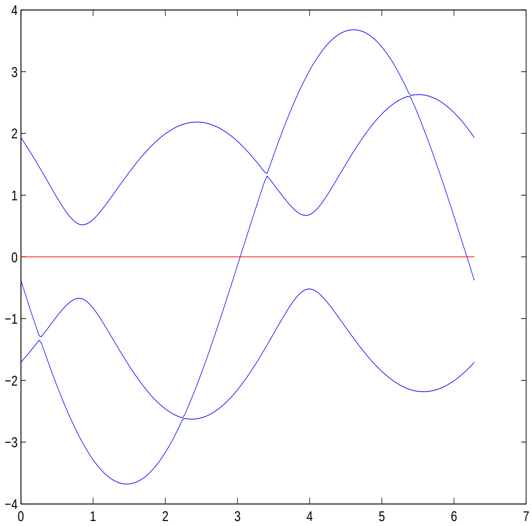

FIG. 4.6. $n = 3$, *eigenvalues of* A_t *for* $t \in [0, 2\pi]$, *matrix (4.1).*

Another interesting example is

$$A = \begin{bmatrix} -0.0786619 & -1.23435 & 0.0558012 \\ -0.681657 & 0.288807 & -0.367874 \\ -1.02455 & -0.429303 & -0.464973 \end{bmatrix}. \tag{4.5}$$

ETNA
Kent State University
http://etna.math.kent.edu

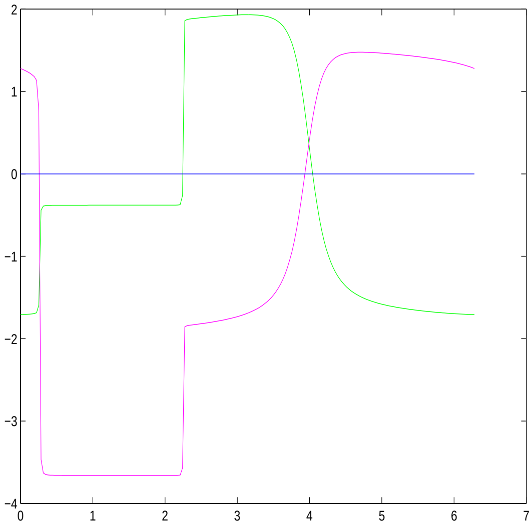

FIG. 4.7. $n = 3$, *values of* $x(t)^T A_1 x(t)$ *(green) and* $x(t)^T A_2 x(t))$ *(magenta), matrix (4.1).*

This matrix has three real eigenvalues and 0 is in the field of values of A and A^2. The eigenvalues of $A_1^{-1} A_2$ are real but nevertheless there are no λ and μ such that $\lambda A_1 + \mu A_2$ is definite. There are 8 real solutions to the polynomial system,

$$\begin{pmatrix} -0.841489 \\ -0.204167 \\ 0.500212 \end{pmatrix}, \begin{pmatrix} 0.841489 \\ 0.204167 \\ -0.500212 \end{pmatrix}, \begin{pmatrix} -0.436377 \\ 0.0701634 \\ 0.897024 \end{pmatrix}, \begin{pmatrix} -0.370335 \\ -0.758645 \\ 0.536012 \end{pmatrix},$$
$$\begin{pmatrix} 0.436377 \\ -0.0701634 \\ -0.897024 \end{pmatrix}, \begin{pmatrix} 0.370335 \\ 0.758645 \\ -0.536012 \end{pmatrix}, \begin{pmatrix} -0.0626104 \\ 0.443589 \\ -0.894041 \end{pmatrix}, \begin{pmatrix} 0.0626104 \\ -0.443589 \\ 0.894041 \end{pmatrix}.$$

The figure corresponding to this problem is 4.8. We see that one red curve intersects the two blue curves. The 4 other solutions are on the other side of the sphere. Figure 4.9 shows the joint field of values which, like in the previous example, has a triangular-like shape. The black circles are given by the real and imaginary parts of the eigenvalues of $A_1 + \imath A_2$. Figure 4.10 displays the eigenvalues of A_t as functions of t. As in the first example the smallest (resp. largest) eigenvalue is always negative (resp. positive). There are no values of λ and μ such that $\lambda A_1 + \mu A_2$ is definite.

Now we consider an example for which 0 is in the field of values of A and A^2 but without any real solutions. Here is such a matrix

$$A = \begin{bmatrix} -0.265607 & 0.986337 & 0.234057 \\ -1.18778 & -0.518635 & 0.0214661 \\ -2.20232 & 0.327368 & -1.00394 \end{bmatrix}. \quad (4.6)$$

The matrix A has complex eigenvalues, but the matrix $A_1^{-1} A_2$ has only real eigenvalues. There is no real solution, as we can see in Figure 4.11, since the blue and red curves do not intersect. This is because there are real values of λ and μ for which $\lambda A_1 + \mu A_2$ is definite as shown in Figure 4.12. The x (resp. y) axis is μ (resp. λ) and there is a magenta (resp. blue) plus sign when the matrix $\lambda A_1 + \mu A_2$ is positive (resp. negative) definite. The boundaries of this cone are given by Theorem 3.7. The straight lines given by some cases in this result are shown in green in Figure 4.12.

ETNA
Kent State University
http://etna.math.kent.edu

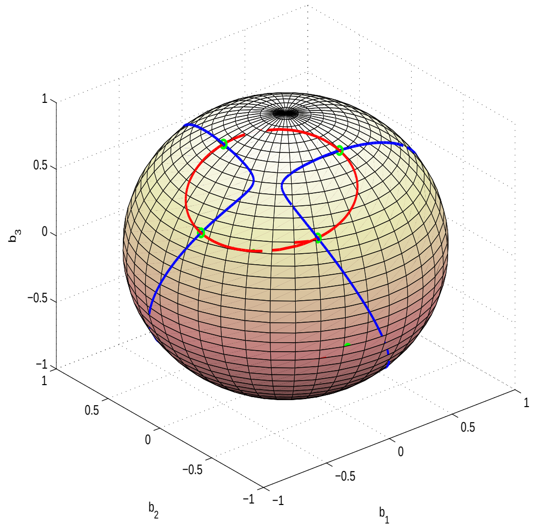

FIG. 4.8. $n = 3$, *the solutions (green circles) in* b_1, b_2, b_3 *frame, matrix (4.5).*

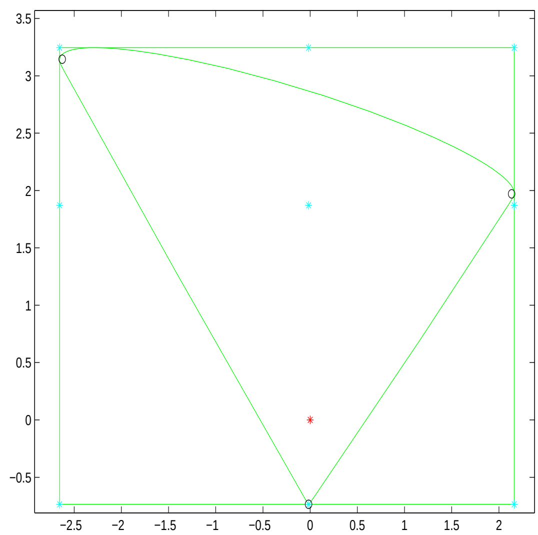

FIG. 4.9. $n = 3$, *joint field of values,* $\{\, (x^T A_1 x, x^T A_2 x),\ x \in \mathbb{R}^3,\ \|x\| = 1\,\}$, *matrix (4.5).*

Figure 4.13 displays the joint field of values. As we can see, $(0, 0)$ is outside this set. Moreover, the ovular shape of the set is quite different from the previous examples. Note that the eigenvalues of $A_1 + \imath A_2$ (black circles) are not on the boundary. Figure 4.14 shows the three eigenvalues of A_t for $t \in [0, 2\pi]$. The circles are the zeros and their colors display the sign of the derivatives. The positive (resp. negative) derivatives are shown in red (resp. black). We see that there are three consecutive red dots. This indicates that there are some values of t such that A_t is positive definite. If we compute the middle point of the interval between the last red circle and the next black one, we obtain a pair $(-0.9176, -0.3974)$ for which $\lambda A_1 + \mu A_2$ is positive definite. The Crawford-Moon algorithm returns the pair $(-0.8974,\ -0.4411)$. As we have seen in Figure 4.12 there is an infinite

ETNA
Kent State University
http://etna.math.kent.edu

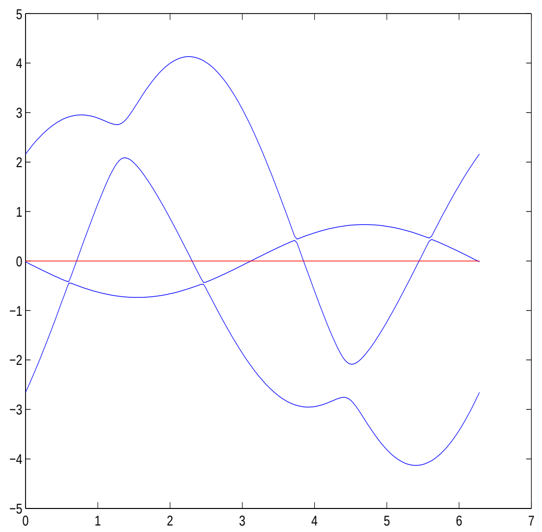

FIG. 4.10. $n = 3$, *eigenvalues of* A_t *for* $t \in [0, 2\pi]$, *matrix (4.5).*

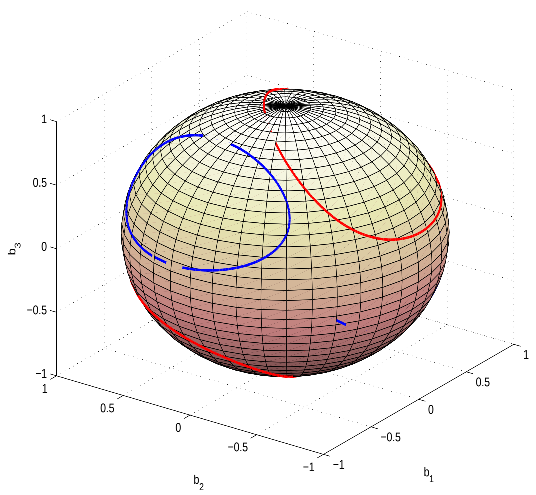

FIG. 4.11. $n = 3$, *no solutions in* b_1, b_2, b_3 *frame, matrix (4.6).*

number of such pairs.

It is interesting to collect some statistics about the number of real solutions for random matrices of order 3 obtained using the Matlab function `randn`. Out of 1500 polynomial systems for real stagnation vectors, 1500 have 8 solutions as they should, but some have only complex solutions. The numbers of real solutions are given in Table 4.1. Remember that the number of real solutions is a multiple of 4. More than one third of the systems do not have a real stagnation vector. There are 69 systems (resp. 228) for which 0 is outside the field of values of A (resp. A^2). This gives at most 297 systems. Therefore, there are many systems for which 0 is in the fields of values of A and A^2 but without real stagnation vector. However, it must also be said that, with random right-hand sides, most of the random systems

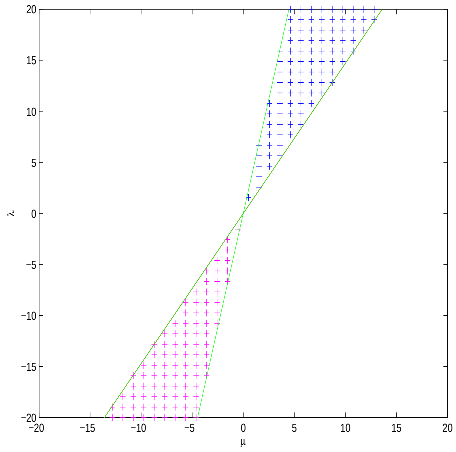

FIG. 4.12. *Definiteness of $\lambda A_1 + \mu A_2$, positive definite (magenta), negative definite (blue), matrix (4.6).*

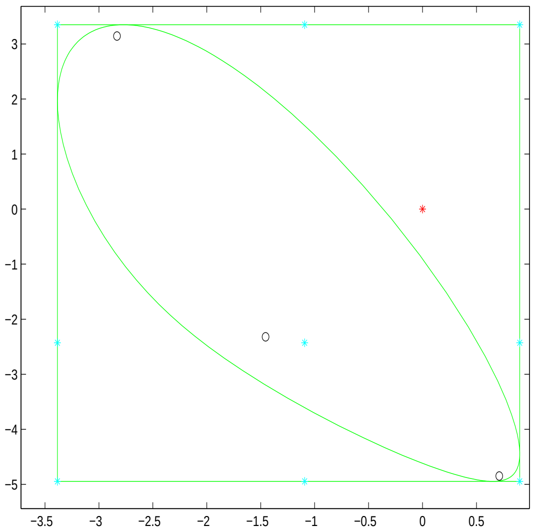

FIG. 4.13. $n = 3$*, joint field of values,* $\{\ (x^T A_1 x, x^T A_2 x),\ x \in \mathbb{R}^3,\ \|x\| = 1\ \}$*, matrix (4.6).*

without stagnation vectors give almost stagnation with a very slow decrease of the residual norm before the last iteration.

TABLE 4.1
Number of real solutions for 1500 random matrices of order 3.

no real sol.	real sol.	4 sol.	8 sol.
603	897	474	423

To conclude with the case $n = 3$, let us consider something else than random matrices, let

$$A = \begin{bmatrix} -1 & 2 & -3 \\ 5 & 4 & -3 \\ 9 & -10 & 1 \end{bmatrix}. \tag{4.7}$$

ETNA
Kent State University
http://etna.math.kent.edu

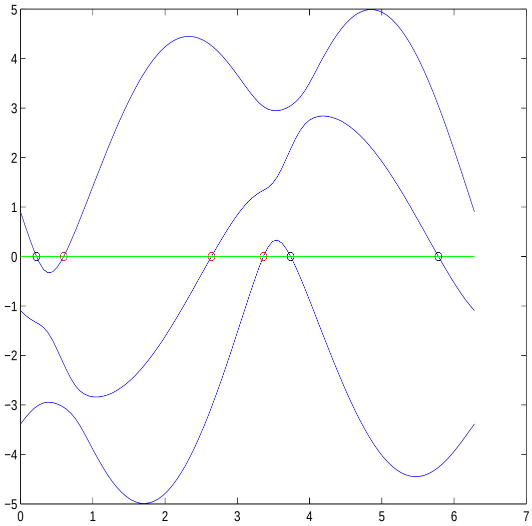

FIG. 4.14. $n = 3$, *eigenvalues of* A_t *for* $t \in [0, 2\pi]$, *matrix (4.6).*

This matrix has one real positive eigenvalue and a pair of complex conjugate eigenvalues. The eigenvalues of $A_1^{-1}A_2$ are real but there are no real λ and μ such that $\lambda A_1 + \mu A_2$ is definite. The apertures of the two cones are large. Each of the red curves intersect the two blue curves. The 8 real solutions to the stagnation system (rounded) are

$$\begin{pmatrix} -0.97358 \\ -0.184462 \\ 0.134598 \end{pmatrix}, \begin{pmatrix} -0.764349 \\ 0.61782 \\ -0.184578 \end{pmatrix}, \begin{pmatrix} 0.97358 \\ 0.184462 \\ -0.134598 \end{pmatrix}, \begin{pmatrix} 0.764349 \\ -0.61782 \\ 0.184578 \end{pmatrix},$$
$$\begin{pmatrix} -0.35124 \\ -0.749866 \\ -0.560652 \end{pmatrix}, \begin{pmatrix} 0.35124 \\ 0.749866 \\ 0.560652 \end{pmatrix}, \begin{pmatrix} -0.00654408 \\ 0.0751576 \\ 0.99715 \end{pmatrix}, \begin{pmatrix} 0.00654408 \\ -0.0751576 \\ -0.99715 \end{pmatrix}.$$

The solutions are displayed in Figure 4.15 and the joint field of values in Figure 4.16. The eigenvalues of $A_1 + \imath A_2$ are (almost) on the boundary. We can check in Figure 4.17 that the matrix A_t is never definite. We can see in Figure 4.18 that the values of $x(t)^T A_1 x(t)$ and $x(t)^T A_2 x(t)$ are either rapidly increasing (or decreasing) or are almost constant. This explains the triangular-like shape of the joint field of values.

4.3. The case $n = 4$. With $n = 4$, there is not much to visualize. However, we can still look at the boundary of the joint field of values of A_1, A_2 and A_3; see an example in Figure 4.19. This is done using a routine of Chi-Kwong Li available on the Web (*http://www.math.wm.edu/∼ckli/*). Points on the boundary are given by the eigenvector corresponding to the largest eigenvalue of a linear combination of A_1, A_2 and A_3. The green box is given by the eigenvalues of A_1, A_2 and A_3. The boundary of the joint field of values has sometimes strange shapes. Figure 4.19 displays an example with 4 real solutions corresponding to the random matrix

$$A = \begin{bmatrix} 1.36526 & -0.310516 & 0.72768 & 0.644051 \\ 2.26211 & 0.42492 & 0.346095 & -0.775557 \\ 0.0979918 & -0.0251637 & -0.563292 & -1.04728 \\ 0.556201 & 0.235534 & 0.0501128 & -0.06832 \end{bmatrix}. \tag{4.8}$$

ETNA
Kent State University
http://etna.math.kent.edu

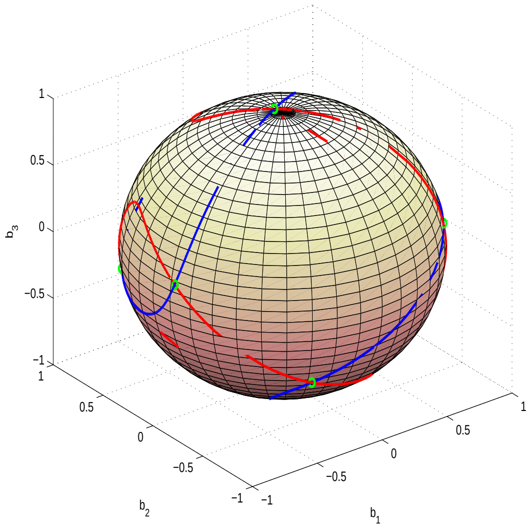

FIG. 4.15. $n = 3$, *the solutions (green circles) in* b_1, b_2, b_3 *frame, matrix* (4.7).

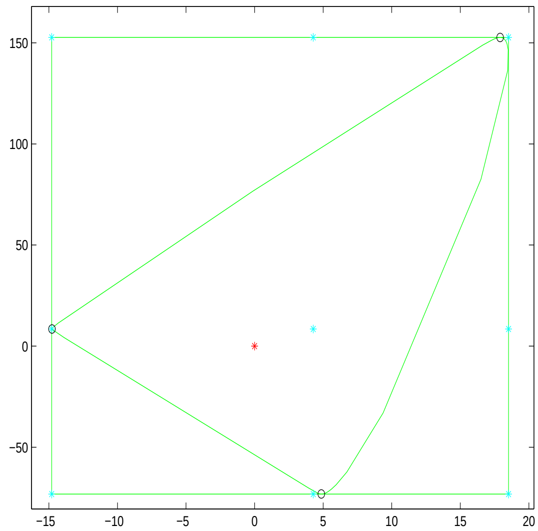

FIG. 4.16. $n = 3$, *joint field of values,* $\{ (x^T A_1 x, x^T A_2 x),\ x \in \mathbb{R}^3,\ \|x\| = 1 \}$, *matrix* (4.7).

This example has two flat portions on the boundary. The red star inside the joint field of values is $(0, 0, 0)$.

The following example has 8 real solutions

$$A = \begin{bmatrix} -0.432565 & -1.14647 & 0.327292 & -0.588317 \\ -1.66558 & 1.19092 & 0.174639 & 2.18319 \\ 0.125332 & 1.18916 & -0.186709 & -0.136396 \\ 0.287676 & -0.0376333 & 0.725791 & 0.113931 \end{bmatrix}.$$

The matrix A has two real eigenvalues and a pair of complex conjugate eigenvalues. An

ETNA
Kent State University
http://etna.math.kent.edu

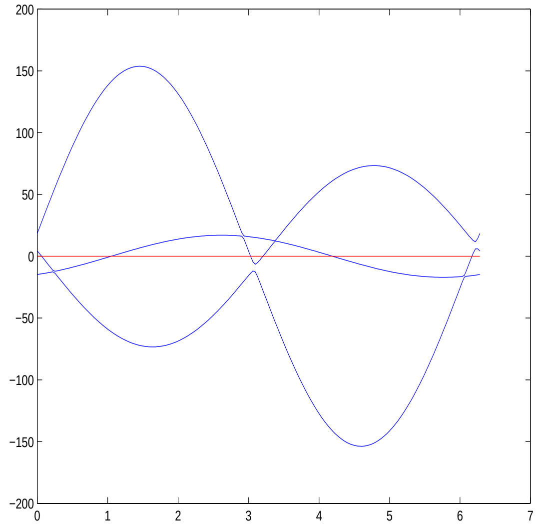

FIG. 4.17. $n = 3$, *eigenvalues of* A_t *for* $t \in [0, 2\pi]$, *matrix (4.7).*

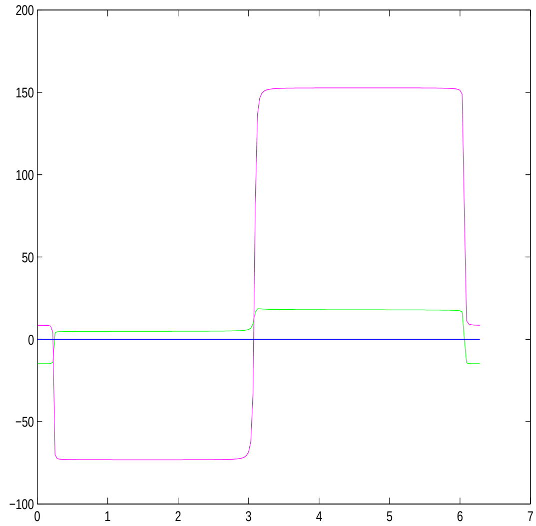

FIG. 4.18. $n = 3$, *values of* $x(t)^T A_1 x(t)$ *(green) and* $x(t)^T A_2 x(t))$ *(magenta), matrix (4.7).*

example without real solutions is

$$A = \begin{bmatrix} 1.06677 & 0.294411 & -0.691776 & -1.44096 \\ 0.0592815 & -1.33618 & 0.857997 & 0.571148 \\ -0.0956484 & 0.714325 & 1.254 & -0.399886 \\ -0.832349 & 1.62356 & -1.59373 & 0.689997 \end{bmatrix}.$$

This matrix has real eigenvalues and the origin is in the fields of values of A_1, A_2 and A_3 but all the solutions to the stagnation system are complex. The point $(0, 0, 0)$ is close but outside the joint field of values. The algorithm PC [52] finds a positive definite linear combination $6.97105\, A_1 + 0.764442\, A_2 - 1.22452\, A_3$ in 28 iterations. The PDC algorithm [58] finds

ETNA
Kent State University
http://etna.math.kent.edu

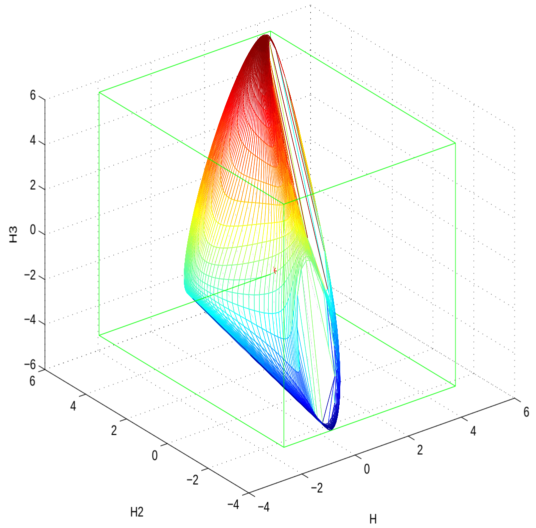

FIG. 4.19. $n = 4$, *boundary of joint field of values,* $\{ (x^T A_1 x, x^T A_2 x, x^T A_3 x),\ x \in \mathbb{R}^4,\ \|x\| = 1 \}$, *matrix (4.8).*

a positive definite linear combination $634.26\, A_1 + 99.2914\, A_2 - 132.612\, A_3$ in 9 iterations with a value of $\sigma = 1000$.

The matrix

$$A = \begin{bmatrix} -1 & 2 & -3 & 1 \\ 5 & 4 & -3 & 2 \\ 9 & -10 & 1 & 3 \\ -7 & 9 & 8 & -10 \end{bmatrix}$$

is an example for which we have 16 real solutions which is the maximum for $n = 4$. In this example, the number of solutions does not seem to be very sensitive to perturbations of the coefficients. The matrix A has two real and a pair of complex conjugate eigenvalues.

For $n = 4$, we also collected statistics on the number of solutions for real random matrices. They are displayed in Table 4.2. Again more than one third of the random matrices do not have a real stagnation vector. For matrices with real solutions there are more cases with 8 solutions than with $4, 12$ or 16 solutions. It seems difficult to identify which characteristics of the matrix A have an influence on the number of real solutions. However, we have seen that this depends on the eigenvalues and eigenvectors of A_1, A_2 and A_3.

TABLE 4.2
Number of real solutions for 1500 random matrices of order 4.

no real sol.	real sol.	4 sol.	8 sol.	12 sol.	16 sol.
597	903	206	523	71	103

4.4. The case $n > 4$. We gathered data about the number of real solutions for random matrices of order 5 and 6. They are displayed in Tables 4.3 and 4.4. The balance between matrices without and with real solutions is more or less 600 to 900. Random matrices with a large number of real solutions are very uncommon.

Let us consider some matrices of order 5 with integer coefficients. For the following matrix, the origin is in the fields of values of A^j, $j = 1, \dots, 5$, which have a circle-like shape

ETNA
Kent State University
http://etna.math.kent.edu

TABLE 4.3
Number of real solutions for 1500 random matrices of order 5.

no real sol.	real sol.	4 s.	8 s.	12 s.	16 s.	20 s.	24 s.	28 s.	32 s.
568	932	159	298	112	291	29	17	9	17

TABLE 4.4
Number of real solutions for 1500 random matrices of order 6.

no real sol.	real sol.	4 s.	8 s.	12 s.	16 s.	20 s.	24 s.	28 s.	32 s.
599	901	96	165	123	268	57	64	32	77

36 s.	40 s.	44 s.	48 s.	52 s.	56 s.	60 s.	64 s.
5	4	2	1	1	1	1	4

and are nested, but there is no real stagnation vector,

$$A = \begin{bmatrix} -10 & -6 & 0 & 8 & 5 \\ 6 & 3 & -3 & 7 & -2 \\ 5 & -10 & 10 & 5 & -3 \\ 16 & 0 & -18 & 0 & -2 \\ 5 & 0 & 4 & 6 & -14 \end{bmatrix}.$$

It is difficult to know if there exists any positive definite linear combination of $A_i, i = 1, \ldots, 4$. With algorithm PDC and a large number of iterations, we found a linear combination for which there is only one negative eigenvalue $-7.8488\ 10^{-14}$. Whether or not this matrix can be considered as being (semi) positive definite is difficult to decide. It may be that the origin is close to the boundary of the joint field of values. The existence (or non-existence) of μ_i such that $A(\mu) = \sum_{i=1}^{4} \mu_i A_i$ is positive definite may eventually be decided by using a symbolic computation package, computing the eigenvalues of $A(\mu)$ as function of the μ_is.

Another example without real solutions is

$$A = \begin{bmatrix} 6 & -20 & -6 & 3 & -10 \\ -6 & -4 & -23 & 9 & -1 \\ 5 & 4 & -12 & -21 & 15 \\ -10 & -3 & 10 & -6 & 0 \\ 0 & 12 & -1 & -7 & 12 \end{bmatrix}.$$

However, for this example, algorithm PDC has found a positive definite linear combination $0.0289275\, A_1 + 0.00253063\, A_2 + 0.000102243\, A_3 + 6.39057\ 10^{-6}\, A_4$ in 80 iterations.

5. Conclusions. We have given a sufficient condition for the non-existence of stagnation vectors for any order n and necessary and sufficient conditions for $n = 3$ and $n = 4$. These conditions have been illustrated with many numerical examples. An open and interesting question is to prove or disprove the converse of the sufficient condition for a general $n > 4$.

Acknowledgments. This paper was started in 2010 during a visit to the Nečas Center of Charles University in Prague supported by a grant of the Jindrich Nečas Center for mathematical modeling, project LC06052 financed by MSMT. The author thanks particularly Zdeněk Strakoš and Miroslav Rozložník for their kind hospitality.

REFERENCES

[1] A. A. ALBERT, *A quadratic form problem in the calculus of variations*, Bull. Amer. Math. Soc., 44 (1938), pp. 250–253.
[2] W. AI, Y. HUANG, AND S. ZHANG, *New results on Hermitian matrix rank-one decomposition*, Math. Program., 128 (2011), pp. 253–283.
[3] E. L. ALLGOWER AND K. GEORG, *Introduction to Numerical Continuation Methods*, SIAM, Philadelphia, 2003.
[4] A. V. ARUTYUNOV, *Positive quadratic forms on intersections of quadrics*, Math. Notes, 71 (2002), pp. 25–33.
[5] W. AUZINGER AND H. J. STETTER, *An elimination algorithm for the computation of all zeros of a system of multivariate polynomial equations*, in Proceedings of the International Conference on Numerical Mathematics, R. P. Agarwal, Y. M. Chow, and S. L. Wilson, eds., International Series of Numerical Mathematics, 86, Birkhäuser, Basel, 1988, pp. 12–30.
[6] ———, *A study of numerical elimination for the solution of multivariate polynomial systems*, Technical Report, Institut f. Angewandte u. Numerische Mathematik, TU Wien, 1989.
[7] L. BRICKMAN, *On the field of values of a matrix*, Proc. Amer. Math. Soc., 12 (1961), pp. 61–66.
[8] E. BROWN AND I. M. SPITKOVSKY, *On flat portions on the boundary of the numerical range*, Linear Algebra Appl., 390 (2004), pp. 75–109.
[9] Q-T. CAI, C-Y. PENG, AND C-S. ZHANG, *Analysis of max-min eigenvalue of constrained linear combinations of symmetric matrices*, in Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE, 2007, pp. 1381–1384.
[10] E. CALABI, *Linear systems of real quadratic forms*, Proc. Amer. Math. Soc., 15 (1964), pp. 844–846.
[11] R. CARDEN, *A simple algorithm for the inverse field of values problem*, Inverse Problems, 25 (2009), 115019, 9 pp.
[12] M-T. CHIEN AND H. NAKAZATO, *Joint numerical range and its generating hypersurface*, Linear Algebra Appl., 432 (2010), pp. 173–179.
[13] C. CHORIANOPOULOS, P. PSARRAKOS, AND F. UHLIG, *A method for the inverse numerical range problem*, Electron. J. Linear Algebra 20 (2010), pp. 198-206.
[14] C. R. CRAWFORD, *Algorithm 646: PDFIND: a routine to find a positive definite linear combination of two real symmetric matrices*. ACM Trans. Math. Software, 12 (1986), pp. 278–282.
[15] C. R. CRAWFORD AND Y. S. MOON, *Finding a positive definite linear combination of two Hermitian matrices*, Linear Algebra Appl., 51 (1983), pp. 37–48.
[16] L. L. DINES, *On the mapping of quadratic forms*, Bull. Amer. Math. Soc., 47 (1941), pp. 494–498.
[17] ———, *On the mapping of n quadratic forms*, Bull. Amer. Math. Soc., 48 (1942), pp. 467–471.
[18] ———, *On linear combinations of quadratic forms*, Bull. Amer. Math. Soc., 49 (1943), pp. 388–339.
[19] W. F. DONOGHUE, *On the numerical range of a bounded operator*, Michigan Math. J., 4 (1957), pp. 261–263.
[20] M. K. H. FAN AND A. L. TITS, *On the generalized numerical range*, Linear and Multilinear Algebra, 21 (1987), pp. 313–320.
[21] P. FINSLER, *Über das Vorkommen definiter und semidefiniter Formen in Scharen quadratischer Formen*, Comment. Math. Helv., 9 (1937), pp. 188–192.
[22] C. B. GARCIA AND T. Y. LI, *On the number of solutions to polynomial systems of equations*, SIAM J. Numer. Anal., 17 (1980), pp. 540–546.
[23] GEOMETRIC TOOLS. http://www.geometrictools.com.
[24] E. GUTKIN, E. A. JONCKHEERE, AND M. KAROW, *Convexity of the joint numerical range: topological and differential geometric view points*, Linear Algebra Appl., 376 (2003), pp. 143–171.
[25] C. HAMBURGER, *Two extensions to Finsler's recurring theorem*, Appl. Math. Optim., 40 (1999), pp. 183–190.
[26] M. HESTENES, *Pairs of quadratic forms*, Linear Algebra and Appl., 1 (1968), pp. 397–407.
[27] M. HESTENES AND E. J. MCSHANE, *A theorem on quadratic forms and its application in the calculus of variation*, Trans. Amer. Math. Soc., 47 (1940), pp. 501–512.
[28] N. HIGHAM, F. TISSEUR, AND P. VAN DOOREN, *Detecting a definite Hermitian pair and a hyperbolic or elliptic quadratic eigenvalue problem, and associated nearness problems*, Linear Algebra Appl., 351/352 (2002), pp. 455–474.
[29] J.-B. HIRIART-URRUTY AND M. TORKI, *Permanently going back and forth between the "quadratic world" and the "convexity world" in optimization*, Appl. Math. Optim., 45 (2002), pp. 169–184.
[30] R. A. HORN AND C. R. JOHNSSON, *Topics in Matrix Analysis*, Cambridge University Press, Cambridge, 1991.
[31] M. HUHTANEN AND O. SEISKARI, *Computational geometry of positive definiteness*, Preprint, Departement of Mathematics and System Analysis, Aalto University, 2011.

[32] K. D. IKRAMOV, *Matrix pencils: theory, applications and numerical methods*, J. Soviet Math., 64 (1993), pp 783–853.
[33] K. JBILOU AND H. SADOK, *Analysis of some extrapolation methods for solving systems of linear equations*, Numer. Math., 70 (1995), pp. 73–89.
[34] C. R. JOHNSON, *Numerical determination of the field of values of a complex matrix*, SIAM J. Numer. Anal., 15 (1978), pp. 595–602.
[35] D. KEELER, L. RODMAN, AND I. M. SPITKOVSKY, *The numerical range of* 3×3 *matrices*, Linear Algebra Appl., 252 (1997), pp. 115–139.
[36] V. L. KLEE, *Separation properties for convex cones*, Proc. Amer. Math. Soc., 6 (1955), pp. 313–316.
[37] C.-K. LI AND L. RODMAN, *Shapes and computer generation of numerical ranges of Krein space operators*, Electron. J. Linear Algebra, 3 (1998), pp. 31–47.
[38] M. MARCUS, *Pencils of real symmetric matrices and the numerical range*, Aequationes Math., 17 (1978), pp. 91–103.
[39] G. MEURANT, *Necessary and sufficient conditions for GMRES complete and partial stagnation*, submitted, (2011).
[40] ———, *The computation of isotropic vectors*, Numer. Algorithms, published online March 14, (2012).
[41] B. T. POLYAK, *Convexity of quadratic transformations and its use in control and optimization*, J. Optim. Theory Appl., 99 (1998), pp. 553–583.
[42] P. PSARRAKOS, *Definite triples of Hermitian matrices and matrix polynomials*, J. Comput. Appl. Math., 151 (2003), pp. 39–58.
[43] P. PSARRAKOS AND M. TSATSOMEROS, *Numerical range: (in) a matrix nutshell, Parts 1 / 2*, Math. Notes Washington State Univ, 45 (2002) / 46 (2003).
[44] W. T. REID, *A theorem on quadratic forms*, Bull. Amer. Math. Soc., 44 (1938), pp. 437-440.
[45] L. RODMAN AND I. M. SPITKOVSKY, 3×3 *matrices with a flat portion on the boundary of the numerical range*, Linear Algebra Appl., 397 (2005), pp. 193–207.
[46] Y. SAAD AND M. H. SCHULTZ, *GMRES: a generalized minimum residual algorithm for solving nonsymmetric linear systems*, SIAM J. Sci. Statist. Comput., 7 (1986), pp. 856–869.
[47] V. SIMONCINI, *On a non-stagnation condition for GMRES and application to saddle point matrices*, Electron. Trans. Numer. Anal., 37 (2010), pp. 202–213.
http://etna.mcs.kent.edu/vol.37.2010/pp202-213.dir/pp202-213.html
[48] V. SIMONCINI AND D. B. SZYLD, *New conditions for non-stagnation of minimal residual methods*, Numer. Math., 109 (2008), pp. 477–487.
[49] A. J. SOMMESE AND C. W. WAMPLER, *The Numerical Solution of Systems of Polynomials Arising in Engineering and Science*, World Scientific, Hackensack, NJ, 2005.
[50] H. J. STETTER, *Numerical Polynomial Algebra*, SIAM, Philadelphia, 2004.
[51] O. TAUSSKY-TODD, *Positive definite matrices*, in Inequalities: proceedings of a symposium held at Wright-Patterson air force base, O. Shisha ed., Academic Press, New York, 1967, pp. 309–319.
[52] L. TONG, Y. INOUYE, AND R-W. LIU, *A finite step global convergence algorithm for the parameter estimation of multichannel MA process*, IEEE Trans. Signal Process., 40 (1992), pp. 2547–2558.
[53] N-K. TSING AND F. UHLIG, *Inertia, numerical range and zeros of quadratic forms for matrix pencils*, SIAM J. Matrix Anal. Appl., 12 (1991), pp. 146–159.
[54] F. UHLIG, *The number of vectors jointly annihilated by two real quadratic forms determines the inertia of matrices in the associated pencil*, Pacific J. Math.,49 (1973), pp. 537–542.
[55] ———, *On the maximal number of linearly independent real vectors annihilated simultaneously by two real quadratic forms*, Pacific J. Math., 49 (1973), pp. 543–560.
[56] ———, *Definite and semidefinite matrices in a real symmetric matrix pencil*, Pacific J. Math., 49 (1973), pp. 561–568.
[57] ———, *An inverse field of values problem*, Inverse Problems, 24 (2008), 055019, 19 pp.
[58] A. ZAIDI, *Positive definite combination of symmetric matrices*, IEEE Trans. Signal Process., 53 (2005), pp. 4412–4416.
[59] I. ZAVORIN, *Analysis of GMRES convergence by spectral factorization of the Krylov matrix*, Ph.D. Thesis, Department of Computer Science, University of Maryland, (2001).
[60] I. ZAVORIN, D. P. O'LEARY, AND H. ELMAN, *Complete stagnation of GMRES*, Linear Algebra Appl., 367 (2003), pp. 165–183.