International Journal of Algebra, Vol. 5, 2011, no. 9, 429 - 449

The Level Curves of the Angle Function of a Positive Definite Symmetric Matrix

Neeraj Bajracharya

Department of Mathematics and Computer Science
Southern Arkansas University
Magnolia, AR 71754, USA
NBajracharya@saumag.edu

Abstract

Given a real $n \times n$ matrix A, write ϕ_A for the maximum angle by which A rotates any unit vector: $\phi_A := \sup_{x \in S^{n-1}} \angle(x, Ax)$. Suppose that A and B are positive definite symmetric (PDS) $n \times n$ matrices. Then their Jordan product $\{A, B\} := AB + BA$ is also symmetric, but not necessarily positive definite. If $\phi_A + \phi_B \geq \frac{\pi}{2}$, then there exists $S \in \mathrm{SO}_n$ such that $\{A, SBS^{-1}\}$ is indefinite. Of course, if A and B commute, then $\{A, B\}$ is positive definite. Our work grows from the following question: if A and B are commuting positive definite symmetric matrices such that $\phi_A + \phi_B \geq \frac{\pi}{2}$, what is $\inf \big\{\phi_S : S \in \mathrm{SO}_n, \{A, SBS^{-1}\}$ indefinite$\big\}$? In this paper we describe the level curves of the angle function $x \mapsto \angle(x, Ax)$ of a 3×3 PDS matrix, and discuss their interaction with those of a second such matrix.

Mathematics Subject Classification: 15A18, 15A99

Keywords: Jordan product, Jordan Algebra, Angle of an operator, Positive Definite Symmetric Matrix, Level curves

1 Introduction

Given a real $n \times n$ matrix A, write ϕ_A for the maximum angle by which A rotates any unit vector:

$$\phi_A := \sup_{x \in S^{n-1}} \angle(x, Ax).$$

Here S^{n-1} is the unit sphere in $\mathbb{R}^n$. If Ax is zero, $\angle(x, Ax)$ is regarded as $\frac{\pi}{2}$. It is not difficult to see that a symmetric matrix A is positive definite if and only if $\phi_A < \frac{\pi}{2}$.

Suppose that A and B are positive definite symmetric (PDS) $n \times n$ matrices. Then their Jordan product $\{A, B\} := AB + BA$ is also symmetric, but not necessarily positive definite. In fact, $\phi_{\{A,B\}} \leq \phi_A + \phi_B$, so if $\phi_A + \phi_B < \frac{\pi}{2}$ then $\{A, B\}$ is positive definite. However, if $\phi_A + \phi_B \geq \frac{\pi}{2}$, then the eigenvalues of A and B alone do not determine whether or not $\{A, B\}$ is positive definite; the relative positions of their eigenvectors also play a role. Speaking roughly, the further the eigenvectors of A are from those of B, the less likely $\{A, B\}$ is to be PDS. Moreover, there is always some "rotation" B_{rot} of B such that $\{A, B_{\text{rot}}\}$ is indefinite. (we use the term indefinite to denote matrices which are neither positive nor negative definite). More precisely, if $\phi_A + \phi_B \geq \frac{\pi}{2}$ then it is possible to find a special orthogonal matrix $S \in \text{SO}_n$ such that $\{A, B^S\}$ is indefinite. Here B^S denotes SBS^{-1}.

When A and B commute, their eigenvectors are the same and $\{A, B\}$ is PDS. This is in some sense the case in which $\{A, B\}$ is the farthest from being indefinite. This paper grew out of the following question: if A and B are commuting PDS matrices such that $\phi_A + \phi_B \geq \frac{\pi}{2}$, what is the "smallest rotation" S such that $\{A, B^S\}$ is indefinite? Of course, for $n \geq 4$ in general S is not a rotation but a product of orthogonal rotations. In order to state the problem precisely, we define

$$\Phi(A, B) := \inf\{\phi_S : S \in \text{SO}_n \text{ and } \{A, B^S\} \text{ is indefinite}\}.$$

Problem 1 *Compute $\Phi(A, B)$ for commuting PDS matrices A and B.*

There is important related work by Strang [3], who deduced bounds on the extremal eigenvalues of $\{A, B\}$ in answer to a research problem proposed by Taussky-Todd [5]. Other related work includes that of Nicholson [2], who (independently of Strang) gave a sufficient condition for $\{A, B\}$ to be positive definite, and Alikakos and Bates [4], who established bounds for all the eigenvalues of $\{A, B\}$. Recently, Conley, Pucci, and Serrin [1] applied Strang's results in establishing the domain of validity of a certain version of the maximum principle.

In this paper we study the level curves of the angle function $x \mapsto \angle(x, Ax)$ of a 3×3 PDS matrix and discuss their interaction with those of a second such matrix. In Section 2 we recall a formula for the maximum angle of rotation of a PDS matrix A, equivalent to Kantorovich's Inequality [6]. In Section 2.1 we review the solution to Problem 1 in the two-dimensional case. In Section 2.2 we solve Problem 1 in three dimensions in the limiting case when A and B are diagonal matrices with $\phi_A + \phi_B = \pi/2$.

In Section 3 we compute the angles of rotation of certain elements of SO_3 pertaining to Problem 1. Fix PDS matrices A and B and vectors a and b. There is a unique element $S_{\max}$ of SO_3 such that $S_{\max} b = a$, and $\angle(Aa, S_{\max} Bb) = \angle(a, Aa) + \angle(b, Bb)$. In Proposition 3.1 we compute its angle of rotation.

When $\angle(a, Aa) + \angle(b, Bb)$ exceeds $\pi/2$, there are exactly two elements $S_\perp^\pm$ of SO_3 such that $S_\perp^\pm b = a$, and $\angle(Aa, S_\perp^\pm Bb) = \pi/2$. In Proposition 3.2 we compute their angles of rotation. Solving Problem 1 amounts to minimizing the formula computed in Proposition 3.2 over a and b.

In Section 4 we analyze the level curves of the angle function of a 3×3 diagonal matrix A with diagonal entries $0 < A_1 \leq A_2 \leq A_3$. For $x = (x_1, x_2, x_3)$ on S^2, define $s = (s_1, s_2, s_3)$ by $s_i = x_i^2$. In s-coordinates, the α-level curve of the angle function $x \mapsto \angle(x, Ax)$ on S^2 is given by $\left(\sum_{i=1}^3 A_i s_i\right)^2 = \left(\sum_{i=1}^3 A_i^2 s_i\right) \cos^2 \alpha$. The image of the sphere in s-coordinates is the standard simplex in the plane $s_1 + s_2 + s_3 = 1$, each point of which corresponds to exactly one point in each octant of S^2. The solutions of the above equation on $s_1 + s_2 + s_3 = 1$ form a parabola, and as α varies the resulting parabolas have parallel axes.

In Section 4.1 we depict the level curves on S^2 and on the standard simplex. They depend qualitatively on which of A_2/A_1 and A_3/A_2 is bigger. In Section 4.2 we prove that, apart from the trivial case when A has repeated eigenvalues, the only case of a circular level curve occurs when $A_2/A_1 = A_3/A_2$, for the angle at which in s-coordinates the curve is doubly tangent to the boundary of the simplex. In Section 4.3 we compute the vertices and axes of the level parabolas in s-coordinates.

Section 5 contains the result which is a step towards minimizing the angles of rotation computed in Section 3. Although the formulas in Section 3 are explicit, we could find only geometric approaches to their minimization.

Fix diagonal PDS matrices A and B with increasing entries, and angles α and β. For α small, the α-level curve of the angle function of A on S^2 consists of six disjoint simple closed curves, one around each of $\pm e_1, \pm e_2$, and $\pm e_3$. Write $\mathcal{C}(A, \alpha)$ for the one around e_1. It is symmetric under reflection through e_2 and e_3, and the points on it closest to e_1 are $a_\pm$, its intersections with the quarter great circles from e_1 to $\pm e_3$.

Assume that β is small in the same sense that α is, and define $b_\pm$ to be the points on $\mathcal{C}(B, \beta)$ closest to e_1. Theorem 5.1 states that the vector a on $\mathcal{C}(A, \alpha)$ such that the angle of rotation of $S_{\max}$ of (a, b_-) is minimal is a_+, the point on $\mathcal{C}(A, \alpha)$ opposite to b_- on the great circle through e_1 and e_3. The proof occupies the rest of the section and depends on the results of Section 4.

We expect that in fact the minimum angle of $S_{\max}$ of (a, b) with (a, b) on $\mathcal{C}(A, \alpha) \times \mathcal{C}(B, \beta)$ occurs at $(a_\pm, b_\mp)$. Indeed, it seems "intuitively obvious" that when A and B are both diagonal with increasing entries, the solution of the Problem 1 should be simply $\Phi\left(\begin{pmatrix} A_1 & \\ & A_3 \end{pmatrix}, \begin{pmatrix} B_1 & \\ & B_3 \end{pmatrix} \right)$.

The real interest in Problem 1 lies in the case that when A and B are simultaneously diagonalized so that A is increasing, then B is not increasing. As we show in Section 2, here the solution is already surprising in the 3×3 case with $\phi_A + \phi_B = \pi/2$. Perhaps a numerical investigation of the 3×3 case with arbitrary $\phi_A + \phi_B$ would be illuminating.

2 Angles of Rotation

The following proposition, which is equivalent to Kantorovich's Inequality [6], gives the formula for the maximum angle of rotation of a PDS matrix in terms of its extremal eigenvalues. We omit the proof.

Proposition 2.1 *Let A be a PDS $n \times n$ matrix with eigenvalues $0 < A_1 \leq A_2 \leq A_3 \leq \cdots \leq A_n$. Then*

$$\cos\phi_A = 2\frac{\sqrt{A_1 A_n}}{A_1 + A_n}, \qquad \sin\phi_A = \frac{A_n - A_1}{A_1 + A_n}.$$

Moreover, $\angle(v, Av) = \phi_A$ if and only if v is a scalar multiple of $(v_1\sqrt{A_n} + v_n\sqrt{A_1})$, where v_1 and v_n are unit eigenvectors of A with eigenvalues A_1 and A_n, respectively.

Let A be a diagonal PDS matrix with diagonal entries $0 < A_1 < A_2 < \cdots < A_n$. Then by Proposition 2.1, there are exactly four unit vectors a such that $\angle(a, Aa) = \phi_A$, namely, $\frac{1}{\sqrt{A_1+A_n}}(\pm\sqrt{A_n}, 0, \cdots, 0, \pm\sqrt{A_1})$. Moreover, these choices lie one on each quadrant of the space spanned by e_1 and e_n. Without loss of generality, we may assume a to be the vector on the positive quadrant. We have the following corollary of Proposition 2.1:

Corollary 2.2 $\angle(a, e_1) = \frac{\pi}{4} - \frac{\phi_A}{2}$.

Let $R_\theta = \left(\begin{smallmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{smallmatrix}\right)$. The next two lemmas are well-known.

Lemma 2.3 *Every $S \in \mathrm{SO}_n$ is orthogonally similar to a block diagonal matrix with diagonal entries $R_{\theta_1}, R_{\theta_2}, \cdots, R_{\theta_{\lfloor n/2\rfloor}}$ if n is even, and $R_{\theta_1}, R_{\theta_2}, \cdots, R_{\theta_{\lfloor n/2\rfloor}}, 1$ if n is odd.*

Lemma 2.4 *For $S \in \mathrm{SO}_n$ let $\theta_1, \theta_2, \cdots, \theta_{\lfloor n/2\rfloor}$ be the orthogonal angles of rotation. Then the maximum angle by which S rotates any vector is $\max_i\{\theta_i\}$.*

Lemma 2.5 *Let A and B be two $n \times n$ PDS matrices and let S be in SO_n. The Jordan product $\{A, B^S\}$ is indefinite if and only if there exist non-zero vectors $a, b \in \mathbb{R}^n$ such that $Sb = a$ and $\angle(Aa, SBb) \geq \frac{\pi}{2}$.*

Proof If $\{A, B^S\}$ is indefinite, then there exists a non-zero $a \in \mathbb{R}^n$ such that $a \cdot \{A, B^S\}a \leq 0$. Since A, B and B^S are all symmetric, $a \cdot AB^S a = Aa \cdot B^S a = a \cdot B^S Aa$, so $a \cdot \{A, B^S\}a = 2(Aa \cdot B^S a)$. Let $b = S^{-1}a$. Then $Aa \cdot SBb \leq 0$, so $\angle(Aa, SBb) \geq \frac{\pi}{2}$.

For the converse, use the same argument in the reverse direction.

2.1 The 2×2 Case

Suppose A and B are 2×2 diagonal PDS matrices. Without loss of generality we may rescale and assume $A = \left(\begin{smallmatrix} 1 & 0 \\ 0 & k \end{smallmatrix}\right)$ and $B = \left(\begin{smallmatrix} 1 & 0 \\ 0 & l \end{smallmatrix}\right)$, where $k = A_2/A_1$ and $l = B_2/B_1$. Let S be in SO_2. Recall that any $S \in \mathrm{SO}_2$ takes the form R_θ, where $\phi_S = \theta$.

We want to find S with the minimum ϕ_S such that the Jordan product $\{A, B^S\} = \left(\begin{smallmatrix} 2(\cos^2\theta + l\sin^2\theta) & (1-l)(1+k)\sin\theta\cos\theta \\ (1-l)(1+k)\sin\theta\cos\theta & 2(k\sin^2\theta + kl\cos^2\theta) \end{smallmatrix}\right)$ is indefinite. It is indefinite if and only if $\mathrm{Det}\{A, B^S\} \leq 0$. Computation gives

$$\mathrm{Det}\{A, B^S\} = 4lk - \frac{1}{4}\big\{(1-l)^2(1-k)^2\big\}(\sin 2\theta)^2. \tag{1}$$

Hence for $\{A, B^S\}$ to be indefinite, we must have $4lk - \frac{1}{4}\big\{(1-l)^2(1-k)^2\big\}(\sin 2\theta)^2 \leq 0$. Thus $\{A, B^S\}$ is indefinite for some S if and only if $16lk \leq (1-l)^2(1-k)^2$.

The following result is proved in [1]:

Proposition 2.6 *The minimum ϕ_S such that $\{A, B^S\}$ is indefinite satisfies*

$$\cos^2\phi_S = \frac{1}{2}\Big(1 + \frac{\sqrt{(l-1)^2(k-1)^2 - 16kl}}{(l-1)(k-1)}\Big).$$

2.2 Some Examples in the 3×3 Case

Definition 2.7 *For $\omega \in S^2$ and $\theta \in \mathbb{R}$, let $R(\omega, \theta)$ be the element of* SO_3 *which rotates by θ counterclockwise around ω.*

Fix matrices $A = \left(\begin{smallmatrix} A_1 & 0 & 0 \\ 0 & A_2 & 0 \\ 0 & 0 & A_3 \end{smallmatrix}\right)$ and $B = \left(\begin{smallmatrix} B_1 & 0 & 0 \\ 0 & B_2 & 0 \\ 0 & 0 & B_3 \end{smallmatrix}\right)$ such that $0 < A_1 < A_2 < A_3$, the B_i are positive and distinct, and $\phi_A + \phi_B = \pi/2$. We will find $\Phi(A, B)$ in this case. By Lemma 2.5, for S orthogonal $\{A, B^S\}$ is indefinite if and only if there exist non-zero vectors a, b such that $Sb = a$ and $\angle(Aa, SBb) \geq \pi/2$. For any a, b, A, B, and S,

$$\angle(Aa, SBb) \leq \angle(Aa, a) + \angle(Sb, SBb) \leq \phi_A + \phi_B.$$

Here $\phi_A + \phi_B = \pi/2$, so $\{A, B^S\}$ is indefinite if and only if there exists a with $\angle(a, Aa) = \phi_A$ and b with $\angle(b, Bb) = \phi_B$ such that $\angle(Aa, SBb) = \pi/2$. By Proposition 3.1 of the next section, there exists exactly one such S for each such pair (a, b). We write $S_{\max}(a, b)$ for this S.

By Proposition 2.1, there are exactly four unit vectors a satisfying $\angle(a, Aa) = \phi_A$, namely, $\frac{1}{\sqrt{A_1+A_3}}(\pm\sqrt{A_3}, 0, \pm\sqrt{A_1})$. Fix $a = \frac{1}{\sqrt{A_1+A_3}}(\sqrt{A_3}, 0, \sqrt{A_1})$. Similarly, there are exactly four unit vectors b that satisfy $\angle(b, Bb) = \phi_B$. We must deduce which of the four choices of b minimizes $\phi_{S_{\max}(a,b)}$. It is easy to see that $\Phi(A, B)$ is the least of the four choices of $\phi_{S_{\max}}(a, b)$. There are six cases.

The case $B_1 < B_2 < B_3$**:** By Proposition 2.1, the four unit vectors b that satisfy $\angle(b, Bb) = \phi_B$ are $\frac{1}{\sqrt{B_1+B_3}}(\pm\sqrt{B_3}, 0, \pm\sqrt{B_1})$. We denote them by b_I, b_{II}, b_{III}, and b_{IV}, the subscript indicating their quadrant in the (e_1, e_3) plane. Note that a, Aa, b_i and Bb_i all lie in the (e_1, e_3) plane. By Corollary 2.2, the $\pi/4$ line bisects both $\angle(a, Aa)$ and $\angle(b_I, Bb_I)$. Moreover, the four b_i form a rectangle in standard orientation centered on 0, wider than it is tall, and the four Bb_i form the reflection of the b_i rectangle over the $\pi/4$ line. The situation is shown in Figure 1 (left), in which for brevity we define $[x] := x/|x|$.

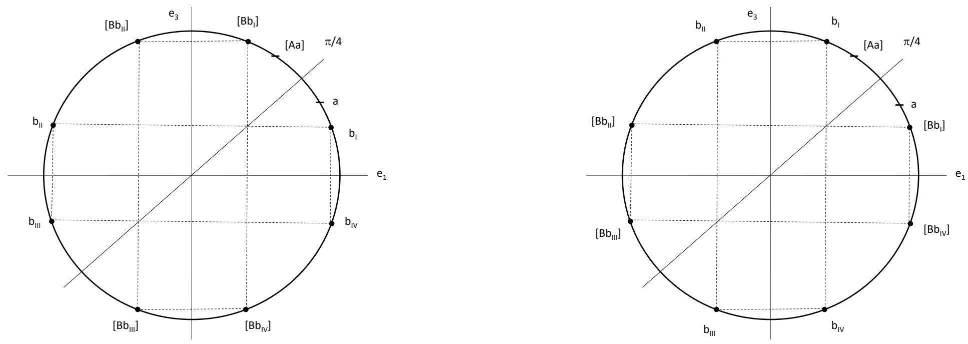


Figure 1: The cases $B_1 < B_2 < B_3$ (on left) and $B_3 < B_2 < B_1$ (on right).

It is easy to see that $\phi_{S(a,b)}$ is minimal for $b = b_{IV}$, and that $S(a, b_{IV})$ is rotation by $\angle(a, b_{VI})$ around e_2. Since $\angle(a, Aa) = \phi_A$ and $\angle(b_i, Bb_i) = \phi_B$, the symmetries of the figure yield $\angle(e_1, a) = \frac{\pi}{4} - \frac{1}{2}\phi_A$ and $\angle(b_{IV}, e_1) = \frac{\pi}{4} - \frac{1}{2}\phi_B$. Therefore since $\phi_A + \phi_B = \pi/2$, here $\Phi(A, B) = \pi/4$.

The case $B_3 < B_2 < B_1$**:** Here the four choices of b_i are given by the same formula as in the last case, but the rectangle they form is taller than it is wide. Thus the situation is as in the preceding case, with b_i and $[Bb_i]$ exchanged (see Figure 1, right).

A similar argument shows that $\phi_{S_{\max}(a,b_i)}$ is minimal for $i = I$, and $S_{\max}(a, b_I)$ is again a rotation by $\pi/4$ about e_2. So here too, $\Phi(A, B) = \pi/4$. (Both of the first two cases are essentially two dimensional.)

The case $B_1 < B_3 < B_2$**:** Here the b_i are $\frac{1}{\sqrt{B_1+B_2}}(\pm\sqrt{B_2}, \pm\sqrt{B_1}, 0)$, on the unit circle in the (e_1, e_2) plane (see Figure 2, left). In contrast to the previous cases, this case is three-dimensional because a and the b_i do not all lie on one plane. Here direct rotation of b_i to a does not give $S_{\max}(a, b_i)$; we must have $[S_{\max}(a, b_i)Bb_i]$ and $[Aa]$ on opposite sides of a on the great circle passing through a and $[Aa]$. Denote this great circle by $\mathrm{GC}(a, Aa)$. $S_{\max}(a, b_i)$ is minimal for b_i equal to either b_I or b_{IV}; we will take $b_i = b_I$. We can factor $S_{\max}(a, b_I)$ as follows. Let

$$S_1 = R(e_3, -\theta_b), \quad S_2 = R(e_1, -\pi/2), \quad S_3 = R(e_2, -\theta_a),$$

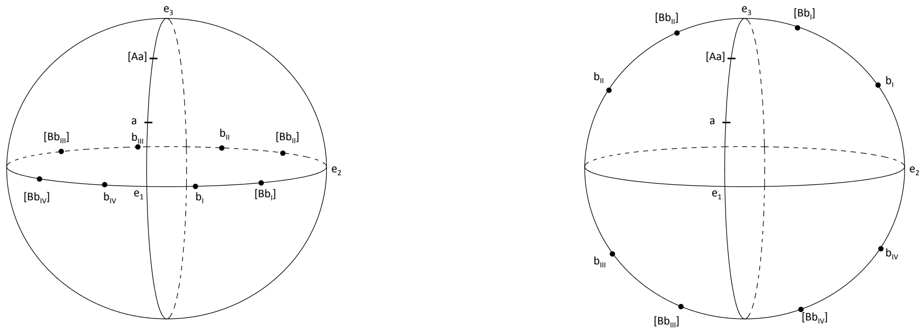


Figure 2: The cases $B_1 < B_3 < B_2$ (on left) and $B_2 < B_1 < B_3$ (on right).

where $\theta_a = \angle(e_1, a)$ and $\theta_b = \angle(e_1, b_I)$. One checks that $S_{\max}(a, b_I) = S_3S_2S_1 = \begin{pmatrix} \cos(\theta_b+\theta_a) & \sin(\theta_b+\theta_a) & 0 \\ 0 & 0 & 1 \\ \sin(\theta_b+\theta_a) & -\cos(\theta_b+\theta_a) & 0 \end{pmatrix}$, so $\mathrm{Tr}(S_{\max}(a, b_I)) = \cos(\theta_b + \theta_a)$. As in the previous cases, Corollary 2.2 gives $\theta_b + \theta_a = \pi/4$, so $\mathrm{Tr}(S_{\max}(a, b_I)) = 1/\sqrt{2}$. Since $\mathrm{Tr}(S) = 1 + 2\cos\phi_S$, $\Phi(A, B) = \pi - \cos^{-1}\left((\sqrt{2} - 1)/2\sqrt{2}\right)$.

The case $B_2 < B_3 < B_1$: Here the four choices of b_i are given by the same formula as in the last case. This argument is similar, but the positions of b_i and $[Bb_i]$ are exchanged. The best choices of b_i are still b_I and b_{IV}. Taking $b_i = b_I$, a similar argument leads to $\mathrm{Tr}(S_{\max}(a, b_I)) = \cos(\theta_a + \theta_b)$. Hence $\Phi(A, B) = \pi - \cos^{-1}\left(\frac{\sqrt{2}-1}{2\sqrt{2}}\right)$.

The case $B_2 < B_1 < B_3$: Here the b_i are $\frac{1}{\sqrt{B_2+B_3}}(0, \pm\sqrt{B_3}, \pm\sqrt{B_2})$, on the unit circle in (e_2, e_3) plane (see Figure 2, right). Since we want $[S_{\max}(a, b_i)Bb_i]$ opposite $[Aa]$ on $\mathrm{GC}(a, Aa)$, $S_{\max}(a, b_i)$ is minimal for b_i equal to either b_{III} or b_{IV}; we will take $b_i = b_{IV}$. We can factor $S_{\max}(a, b_{IV})$ as follows. Let S_1 rotate around e_3 by $-\pi/2$ so as to put b_{IV} on $\mathrm{GC}(e_1, e_3)$, and let S_2 rotate S_1b_{IV} to a: $S_1 = R(e_3, -\pi/2), S_2 = R(e_2, -\theta)$, where $\theta = \angle(a, e_1) + \angle(e_1, b)$. Thus $S_{\max}(a, b_{IV}) = S_2S_1 = \begin{pmatrix} 0 & \cos\theta & -\sin\theta \\ -1 & 0 & 0 \\ 0 & \sin\theta & \cos\theta \end{pmatrix}$, so $\mathrm{Tr}(S_{\max}(a, b_{IV})) = \cos\theta$. Again, $\theta = \theta_a + \theta_b = \frac{\pi}{4}$, so $\Phi(A, B) = \pi - \cos^{-1}\left(\frac{\sqrt{2}-1}{2\sqrt{2}}\right)$.

The case $B_3 < B_1 < B_2$: Here the four choices of b_i are given by the same formula as in the last case, but the positions of b and $[Bb_i]$ are exchanged. Since we want $[S_{\max}(a, b_i)Bb_i]$ and $[Aa]$ on opposite sides of a on $\mathrm{GC}(a, Aa)$, $S_{\max}(a, b_i)$ is minimal for b_i equal to either b_I or b_{II}. Taking $b_i = b_I$, an argument as above again gives $\Phi(A, B) = \pi - \cos^{-1}\left(\frac{\sqrt{2}-1}{2\sqrt{2}}\right)$.

3 The Traces of Some Orthogonal Matrices

For $S \in \mathrm{SO}_3$, $\mathrm{Tr}(S) = 1 + 2\cos\phi_S$, where ϕ_S is the maximum angle by which S rotates any non-zero vector. Suppose that A and B are PDS matrices. Fix arbitrary $a, b \in S^2$, and define $\alpha := \angle(a, Aa)$, and $\beta := \angle(b, Bb)$.

Proposition 3.1 *There exists a unique matrix* $S_{\max}(a,b) \in \mathrm{SO}_3$ *such that* $S_{\max}(a,b)b = a$ *and* $\angle(Aa, S_{\max}(a,b)Bb) = \alpha + \beta$. *Its trace is given by* $\mathrm{Tr}\left(S_{\max}(a,b)\right)$

$$= a \cdot b + \cot\alpha \cot\beta \Big(\frac{a \cdot Bb}{b \cdot Bb} + \frac{b \cdot Aa}{a \cdot Aa} - \frac{Aa \cdot Bb + (b \times Bb) \cdot (a \times Aa)}{(a \cdot Aa)(b \cdot Bb)} - a \cdot b \Big).$$

Proof Consider the orthonormal bases $\{v_1, v_2, v_3\}$ and $\{w_1, w_2, w_3\}$ of $\mathbb{R}^3$ defined by

$$v_1 = b, \quad v_3 = \frac{b \times Bb}{|b \times Bb|}, \quad v_2 = v_3 \times v_1,$$
$$w_1 = a, \quad w_3 = \frac{a \times Aa}{|a \times Aa|}, \quad w_2 = w_3 \times w_1.$$

Since we want to find an orthogonal matrix S such that $Sb = a$ and $\angle(Aa, SBb) = \alpha + \beta$, SBb must lie on the fourth quadrant of the (w_1, w_2) plane, so $Sv_2 = -w_2$ and $Sv_3 = -w_3$. Hence, $\mathrm{Tr}(S) = \sum_{i=1}^{3} v_i \cdot Sv_i$ gives the desired result. Finally, since Sv_i is determined for $1 \le i \le 3$, S is unique.

Proposition 3.2 *Maintain* a, b, α *and* β *as in Proposition 3.1, and assume that* $\alpha + \beta > \pi/2$. *Then there exist exactly two matrices* $S \in \mathrm{SO}_3$ *such that* $Sb = a$ *and* $\angle(Aa, SBb) = \pi/2$. *Their traces are given by* $\mathrm{Tr}(S) =$

$$a \cdot b + \cot^2\alpha \cot^2\beta \Big\{ \Big(\frac{a \cdot Bb}{b \cdot Bb} + \frac{b \cdot Aa}{a \cdot Aa} - \frac{Aa \cdot Bb + (b \times Bb) \cdot (a \times Aa)}{(a \cdot Aa)(b \cdot Bb)} - a \cdot b \Big) \pm \Delta \Big| \big(a - \frac{Aa}{a \cdot Aa}\big) \cdot \big(\frac{b \times Bb}{b \cdot Bb}\big) - \big(b - \frac{Bb}{b \cdot Bb}\big) \cdot \big(\frac{a \times Aa}{a \cdot Aa}\big) \Big| \Big\},$$

where $\Delta = \sqrt{\tan^2\alpha \tan^2\beta - 1}$. *(Note that* $\tan^2\alpha\tan^2\beta \ge 1$ *precisely when* $\alpha + \beta \ge \pi/2$.*)*

Proof Define orthonormal bases $\{v_1, v_2, v_3\}$ and $\{w_1, w_2, w_3\}$ of $\mathbb{R}^3$ as in the last proof. In this case, we must have Sv_2 in the span $\{w_2, w_3\}$. A similar calculation gives the formula for $\mathrm{Tr}(S)$ as desired.

We denote the two choices of S given by $\pm\Delta$ in Proposition 3.2 by $S_\perp^{\pm}(a,b)$, respectively. Define $S_\perp(a,b) := S_\perp^{+}(a,b)$. Note that in the limiting case $\Delta = 0, \alpha + \beta = \frac{\pi}{2}$ and $S_\perp^{+}(a,b) = S_\perp^{-}(a,b) = S_{\max}(a,b)$.

Proposition 3.3 *For* $\phi_A + \phi_B \ge \frac{\pi}{2}$, $\Phi(A,B) = \min\{\phi_{S_\perp(a,b)} : \alpha + \beta \ge \frac{\pi}{2}\}$.

Proof Clearly the left side is less than or equal to the right. To prove the converse, suppose that $\{A, B^S\}$ is indefinite. Then by Lemma 2.5, $Sb_1 = a_1$ and $\angle(Aa_1, SBb_1) \geq \pi/2$ for some (a_1, b_1), so $\angle(Aa_1, B^S a_1) \geq \pi/2$. Since both A and B^S are PDS, if a_2 is an eigenvector of A we have $\angle(Aa_2, SBS^{-1}a_2) < \pi/2$. Thus by the connectedness of S^2, there exists a_3 such that $\angle(Aa_3, B^S a_3) = \pi/2$, so S is one of $S_\perp^\pm(a_3, S^{-1}a_3)$. Since $\mathrm{Tr}(S_\perp^+(a,b)) \geq \mathrm{Tr}(S_\perp^-(a,b))$, we have $\phi_S \geq \phi_{S_\perp(a_3, S^{-1}a_3)}$. The result follows.

Therefore in three dimensions, Problem 1 amounts to maximizing the function $\mathrm{Tr}(S_\perp(a,b))$ given in Proposition 3.2 over all $a, b \in S^2$ with $\alpha + \beta \geq \pi/2$. It would be interesting to prove the following conjecture.

Conjecture 3.4 *The minimum of $\phi_{S_\perp(a,b)}$ over all (a,b) with $\alpha + \beta \geq \frac{\pi}{2}$ occurs for some (a,b) with $\alpha + \beta = \frac{\pi}{2}$.*

4 The Level Curves of the Angle Function

Now we discuss the level curves of the angle function $x \mapsto \angle(x, Ax)$ of A on the unit sphere in $\mathbb{R}^3$ for a PDS matrix A. We may assume without loss of generality that $A = \begin{pmatrix} A_1 & 0 & 0 \\ 0 & A_2 & 0 \\ 0 & 0 & A_3 \end{pmatrix}$, where $0 < A_1 \leq A_2 \leq A_3$. Recall that ϕ_A is the maximum angle by which the matrix A rotates any non-zero vector x. Thus $\angle(x, Ax) \leq \phi_A$ for all $x \in \mathbb{R}^3$.

Since A is positive definite, $\phi_A < \frac{\pi}{2}$. It follows that the level curves of $x \mapsto \angle(x, Ax)$ are the same as those of

$$F(x) := \cos^2\big(\angle(x, Ax)\big) = \frac{(\sum_i A_i x_i^2)^2}{(\sum_i x_i^2)(\sum_i A_i^2 x_i^2)}. \tag{2}$$

Let us denote the level curve $F = \gamma$ by $\mathcal{L}(A, \gamma)$. Writing s_i for x_i^2, this curve has equation

$$\Big(\sum_i A_i s_i\Big)^2 - \gamma\Big(\sum_i s_i\Big)\Big(\sum_i A_i^2 s_i\Big) = 0. \tag{3}$$

Since $|x| = 1$, we must have $s_1 + s_2 + s_3 = 1$ and $s_i \geq 0$ for all i. Thus we seek solutions of (3) on $\mathcal{T} := \big\{s \in \mathbb{R}^3 : s_1 + s_2 + s_3 = 1, s_i \geq 0\big\}$, which is contained in the plane $\mathcal{P} := \{s : s_1 + s_2 + s_3 = 1\}$. We may write (3) as $s^T M_A(\gamma) s = 0$, where $M_A(\gamma) = \begin{pmatrix} A_1^2(1-\gamma) & A_1A_2 - \frac{\gamma}{2}(A_1^2 + A_2^2) & A_1A_3 - \frac{\gamma}{2}(A_1^2 + A_3^2) \\ A_1A_2 - \frac{\gamma}{2}(A_1^2 + A_2^2) & A_2^2(1-\gamma) & A_2A_3 - \frac{\gamma}{2}(A_2^2 + A_3^2) \\ A_1A_3 - \frac{\gamma}{2}(A_1^2 + A_3^2) & A_2A_3 - \frac{\gamma}{2}(A_2^2 + A_3^2) & A_3^2(1-\gamma) \end{pmatrix}$.

We wish to find all $s \in \mathcal{T}$ such that $s^T M_A(\gamma) s = 0$. Let λ_1, λ_2, and λ_3 be the eigenvalues of $M_A(\gamma)$, and let $\{u_1, u_2, u_3\}$ be an orthonormal eigenbasis corresponding to these eigenvalues, respectively. Thus for some scalars α_1, α_2

and α_3, we can write $s = u_1\alpha_1 + u_2\alpha_2 + u_3\alpha_3$. This leads to $s^T M_A(\gamma)s = \alpha_1^2\lambda_1 + \alpha_2^2\lambda_2 + \alpha_3^2\lambda_3$.

When the λ_i are all positive or all negative, the only solution of $s^T M_A(\gamma)s = 0$ is $s = 0$. If the determinant of $M_A(\gamma)$ is not zero, then $M_A(\gamma)$ has exactly one negative eigenvalue. (This can be shown using a certain Vandermonde-type factorization of $M_A(\gamma)$, but we omit the proof as we do not need the result.) Therefore $s^T M_A(\gamma)s = 0$ defines a cone $\mathcal{M}(A, \gamma)$, and the solution of (3) is the conic $\mathcal{P} \cap \mathcal{M}(A, \gamma)$. At the end of this section we will see that this conic is in fact a parabola for all values of γ, and obtain a formula for said parabola. We graph $\mathcal{T} \cap \mathcal{M}(A, \gamma)$ for some A and γ, and map the result back to S^2 to obtain $\mathcal{L}(A, \gamma)$. Note that each point in $\mathcal{T}$ corresponds to exactly one point in each octant of S^2.

4.1 The Level Curves

PDS matrices fall into several categories, with qualitatively different level curves. For each case, we will give pictures of the level curves on both $\mathcal{T}$ and on S^2 for certain values of γ.

Recall Proposition 2.1, and write $\phi(C_1, C_2)$ for ϕ_C when $C = \begin{pmatrix} C_1 & \\ & C_2 \end{pmatrix}$. The level curves change qualitatively at the γ-values $\cos^2\phi(A_1, A_2), \cos^2\phi(A_2, A_3)$, and $\cos^2\phi(A_1, A_3)$ (note that $\phi(A_1, A_3)$ is ϕ_A). We display the level curves at each of these values, as well as at some intermediate values. When $\gamma = 1$, the parabola touches all three vertices of $\mathcal{T}$ and the level curve on S^2 consists of the six points $\pm e_1, \pm e_2$ and $\pm e_3$.

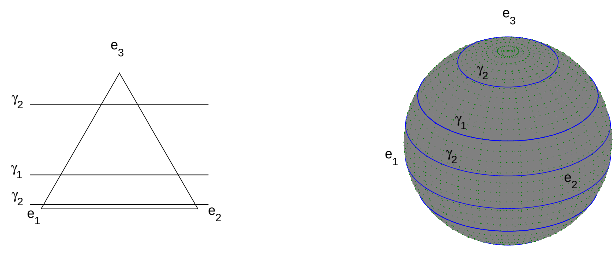


Figure 3: $A = \begin{pmatrix} 1 & & \\ & 1 & \\ & & 3 \end{pmatrix}, \gamma_1 = \frac{3}{4} = \cos^2\phi_A < \gamma_2$

The case $A_1 = A_2 < A_3$. Here the parabola $\mathcal{P} \cap \mathcal{M}(A, \gamma)$ is degenerate and is either two parallel lines or a single "double" line. The level curves have a transition only at $\gamma = \cos^2\phi_A$. For $\gamma < \cos^2\phi_A$, $\mathcal{T} \cap \mathcal{M}(A, \gamma)$ and $\mathcal{L}(A, \gamma)$ are empty. For $\gamma = \cos^2\phi_A$, $\mathcal{T} \cap \mathcal{M}(A, \gamma)$ is a line segment, and $\mathcal{L}(A, \gamma)$ is a pair of opposite latitude lines with respect to the poles $\pm e_3$, one in the northern and one in the southern hemisphere. For $\gamma > \cos^2 \phi_A$, $\mathcal{T} \cap \mathcal{M}(A, \gamma)$ consists of two parallel line segments, and $\mathcal{L}(A, \gamma)$ is two pairs of latitude lines. For $\gamma = 1$, $\mathcal{T} \cap \mathcal{M}(A, \gamma)$ consists of e_3 and the line segment $\overline{e_1 e_2}$, and $\mathcal{L}(A, \gamma)$ is the poles and the equator. We illustrate the situation for $A = \begin{pmatrix} 1 & & \\ & 1 & \\ & & 3 \end{pmatrix}$ in Figure 3.

The case $A_1 < A_2 = A_3$. This is similar to the previous case. Here too the level curves have a transition only at $\gamma = \cos^2 \phi_A$. The only difference is that $\mathcal{L}(A, \gamma)$ consists of latitude lines with $\pm e_1$ as the poles, as $\mathcal{T} \cap \mathcal{M}(A, \gamma)$ consists of line segments parallel to $\overline{e_2 e_3}$. Figures are omitted since they are as in the previous case with e_1 and e_3 exchanged.

The case $1 < \frac{A_2}{A_1} < \frac{A_3}{A_2}$. Here $\phi(A_1, A_2) < \phi(A_2, A_3) < \phi(A_1, A_3) = \phi_A$. We present the level curves for these angles and some intermediate values. Take as an example $A = \begin{pmatrix} 1 & & \\ & 2 & \\ & & 5 \end{pmatrix}$. By Proposition 2.1, $\phi(1, 2) \approx 20°, \phi(2, 5) \approx 25°$, and $\phi(1, 5) = \phi_A \approx 42°$. The corresponding γ values are $\frac{8}{9}, \frac{40}{49}$, and $\frac{5}{9}$ respectively. For $\gamma < \frac{5}{9}$, $\mathcal{L}(A, \gamma)$ is empty. At $\gamma = \frac{5}{9}$, the parabola is tangent to $\overline{e_1 e_3}$, see the γ_1 curve in Figure 4. Thus the level curve on the sphere at this stage consists of four points, a point on each of the four quadrants of the great circle $\mathrm{GC}(e_1, e_3)$.

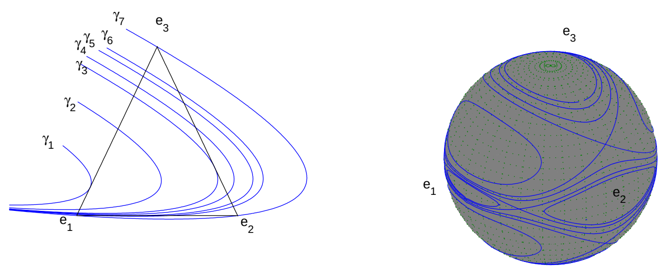


Figure 4: $A = \begin{pmatrix} 1 & & \\ & 2 & \\ & & 5 \end{pmatrix}, \gamma_1 = \frac{5}{9} < \gamma_2 < \frac{40}{49} = \gamma_3 < \gamma_4 < \frac{8}{9} = \gamma_5 < \gamma_6 < 1.$

For $\frac{5}{9} < \gamma < \frac{40}{49}$, the parabola intersects $\overline{e_1 e_3}$ but not $\overline{e_2 e_3}$ or $\overline{e_1 e_2}$. Hence $\mathcal{L}(A, \gamma)$ consists of four roughly oval curves; see the γ_2 curve in Figure 4. At $\gamma = \frac{40}{49}$ the parabola is tangent to $\overline{e_2 e_3}$. The corresponding four points on $\mathrm{GC}(e_2, e_3)$ are saddle points of the angle function of A; see the γ_3 curve in Figure 4. For $\frac{40}{49} < \gamma < \frac{8}{9}$, the parabola intersects $\overline{e_1 e_2}$ and $\overline{e_2 e_3}$, but not $\overline{e_1 e_2}$; see the γ_4 curve in Figure 4. At $\gamma = \frac{8}{9}$, the parabola is tangent to $\overline{e_1 e_2}$, giving the four saddle points of the angle function on $\mathrm{GC}(e_1, e_2)$; see the γ_5 curve in Figure 4.

For $\frac{8}{9} < \gamma < 1$, the parabola intersects all three sides of $\mathcal{T}$ in two points. Hence $\mathcal{L}(A, \gamma)$ consists of six roughly oval curves, around $\pm e_1, \pm e_2$ and $\pm e_3$; see the γ_6 curve in Figure 4. Finally, at $\gamma = 1$ the parabola touches all three vertices of the triangle; see the γ_7 curve in Figure 4.

We remark that the angle function of A on S^2 is minimal at $\pm e_1, \pm e_2$ and $\pm e_3$, maximal at the four points on $\mathrm{GC}(e_1, e_3)$ with $\gamma = \cos^2 \phi(A_1, A_2)$, and has saddles at the eight points on $\mathrm{GC}(e_1, e_2)$ and $\mathrm{GC}(e_2, e_3)$ corresponding to tangency of the parabola with the triangle. It follows from the proof of Proposition 2.1 that it has no other critical points.

The case $1 < \frac{A_3}{A_2} < \frac{A_2}{A_1}$. Here we have $\phi(A_2, A_3) < \phi(A_1, A_2) < \phi(A_1, A_3) = \phi_A$. We give some level curves for $A = \begin{pmatrix} 1 & & \\ & 2 & \\ & & 3 \end{pmatrix}$. By Proposition 2.1, $\phi(2,3) \approx 12°, \phi(1,2) \approx 20°$, and $\phi(1,3) = \phi_A = 30°$. The corresponding γ values are $\frac{24}{25}, \frac{8}{9}$, and $\frac{3}{4}$ respectively. For $\gamma < \frac{3}{4}$, there is no level curve. At $\gamma = \frac{3}{4}$, the parabola is tangent to $\overline{e_1 e_3}$ and so $\mathcal{L}(A, \gamma)$ consists of four points, a point on each of the four quadrants of $\mathrm{GC}(e_1, e_3)$. For $\frac{3}{4} < \gamma < \frac{8}{9}$, the situation is similar to the previous case with γ_1.

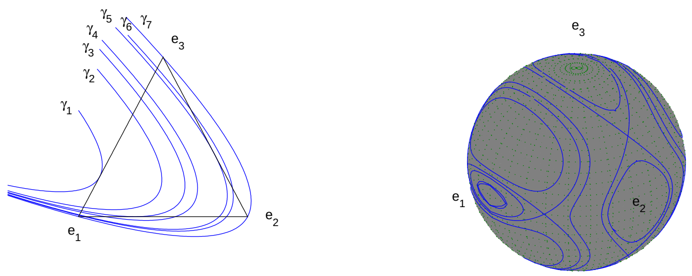


Figure 5: $A = \begin{pmatrix} 1 & & \\ & 2 & \\ & & 3 \end{pmatrix}, \gamma_1 = \frac{3}{4} < \gamma_2 < \frac{8}{9} = \gamma_3 < \gamma_4 < \frac{24}{25} = \gamma_5 < \gamma_6 < 1.$

At $\gamma = \frac{8}{9}$, the parabola is tangent to $\overline{e_1 e_2}$ and the angle function has saddle points on $\mathrm{GC}(e_1, e_2)$; see the γ_3 curve in Figure 5. For $\frac{8}{9} < \gamma < \frac{24}{25}$, the parabola intersects $\overline{e_1 e_2}$ in two points but does not intersect $\overline{e_2 e_3}$; see the γ_4 curve in Figure 5. At $\gamma = \frac{24}{25}$, the parabola is tangent to $\overline{e_2 e_3}$, yielding the saddle points of the angle function on $\mathrm{GC}(e_2, e_3)$; see the γ_5 curve in Figure 5. For $\frac{24}{25} < \gamma < 1$, the parabola intersects all three sides of the triangle, see Figure 5.

The case $1 < \frac{A_2}{A_1} = \frac{A_3}{A_2}$**.** Here $\phi(A_1, A_2) = \phi(A_2, A_3) < \phi(A_1, A_3) = \phi_A$. Consider as an example the matrix $A = \begin{pmatrix} 1 & & \\ & 2 & \\ & & 4 \end{pmatrix}$. By Proposition 2.1, $\phi(1,2) = \phi(2,4) \approx 20°$ and $\phi(1,4) = \phi_A \approx 37°$. The corresponding γ values are $\frac{8}{9}$ and

$\frac{16}{25}$. For $\gamma < \frac{16}{25}$, there are no level curves. For $\gamma = \frac{16}{25}$, the parabola is tangent to $\overline{e_1 e_3}$; see the γ_1 curve in Figure 6. For $\frac{16}{25} < \gamma < \frac{8}{9}$, the parabola intersects only the side $\overline{e_1 e_3}$; see the γ_2 curve in Figure 6.

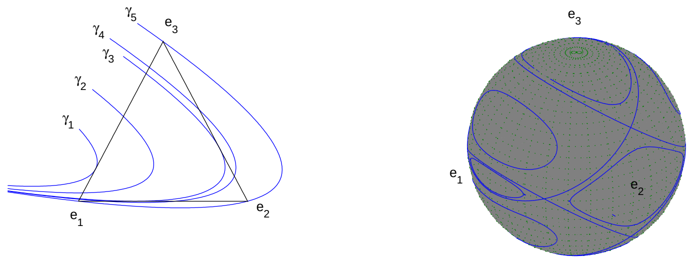


Figure 6: $A = \begin{pmatrix} 1 & & \\ & 2 & \\ & & 4 \end{pmatrix}$, $\gamma_1 = \frac{16}{25} < \gamma_2 < \frac{8}{9} = \gamma_3 < \gamma_4 < 1$.

At $\gamma = \frac{8}{9}$, the parabola is tangent to both $\overline{e_1 e_2}$ and $\overline{e_2 e_3}$, and so $\mathcal{L}(A, \gamma)$ contains all eight saddle points of the angle function; see the γ_3 curve on Figure 6. For $\frac{8}{9} < \gamma < 1$, the situation is as γ_5 and γ_6 curves in Figures 4 curve in 4, respectively.

4.2 Circular Level Curves

Note that the doubly tangent parabola on $\mathcal{T}$ appears to give circular contours on the sphere; see the γ_3 curve in Figure 6. The following proposition confirms this and shows that it is the unique non-trivial occurrence of circular contours.

Proposition 4.1 *The level curve $\mathcal{L}(A, \gamma)$ is a union of circles if and only if either the ratios $\frac{A_2}{A_1}$ and $\frac{A_3}{A_2}$ are equal and $\gamma = \cos^2 \phi(A_1, A_2) = \cos^2 \phi(A_2, A_3)$, or A has repeated eigenvalues and $\gamma \geq \cos^2 \phi_A$.*

Proof Suppose that $\mathcal{L}(A, \gamma)$ contains a circle. If n is any vector normal to this circle, then the equation of the circle (on S^2) is $x \cdot n = N$ for some constant N. Note that since A is diagonal, $x \cdot Mn = N$ must also be contained in $\mathcal{L}(A, \gamma)$ for any M in $\begin{pmatrix} \pm 1 & & \\ & \pm 1 & \\ & & \pm 1 \end{pmatrix} \cong \mathbb{Z}_2^3$. It follows that the product of all $x \cdot Mn = N$ with $M \in \mathbb{Z}_2^3$ and Mn distinct must divide the equation of $\mathcal{L}(A, \gamma)$. Since $\mathcal{L}(A, \gamma)$ is quartic in x, this implies that at least one entry of n is zero. There are two cases.

Suppose first that two entries of n are zero. In this case the circles are centered on $\pm e_1$, $\pm e_2$ or $\pm e_3$. Thus they correspond to line segments parallel to one side of the triangle $\mathcal{T}$ in s-space. This occurs only when A has repeated eigenvalues, in which case the level curve on $\mathcal{T}$ consists of a single line if $\gamma = \cos^2 \phi_A$ and a pair of lines if $\gamma > \cos^2 \phi_A$. On the sphere it consists of two circles if $\gamma = \cos^2 \phi_A$ and four circles if $\gamma > \cos^2 \phi_A$.

Now suppose that only one entry of n is zero. Here by symmetry $\mathcal{L}(A, \gamma)$ contains four congruent circles, so since it is quartic it must be exactly the union of these circles.

Assume first that the second entry of n is zero. Then the normals of the four circles are $n = (\pm n_1, 0, \pm n_3)$, so the equation of $\mathcal{L}(A, \gamma)$ is

$$0 = \prod_{\epsilon_1, \epsilon_3 = \pm 1} (x_1 \epsilon_1 n_1 + x_3 \epsilon_3 n_3 - N) = \left(s_1 n_1^2 + s_3 n_3^2 - N^2\right)^2 - 4 s_1 s_3 n_1^2 n_2^2,$$

where $s_i = x_i^2$ for all i. Expanding and simplifying, we obtain

$$N^4 - 2N^2 n_1^2 s_1 - 2N^2 n_3^2 s_3 + (s_1 n_1^2 - s_3 n_3^2)^2 = 0. \tag{4}$$

Set $k = A_2/A_1$ and $l = A_3/A_2$. We may assume that $A = \begin{pmatrix} 1 & & \\ & k & \\ & & l \end{pmatrix}$. Recall that $\mathcal{L}(A, \gamma)$ is $(s_1 + k s_2 + l s_3)^2 - \gamma(s_1 + k^2 s_2 + l^2 s_3) = 0$, where $s_2 = 1 - s_1 - s_3$. Expanding and simplifying gives

$$\begin{aligned} &k^2(1-\gamma) - \big(\gamma(1+k) - 2k\big)(1-k)s_1 \\ &\quad - \big(2k - \gamma(k+l)\big)(k-l)s_3 + \big((1-k)s_1 - (k-l)s_3\big)^2 = 0. \end{aligned} \tag{5}$$

Since (4) and (5) have the same solutions, they are proportional. We may replace (n, N) by (cn, cN) for any scalar c, so we may assume that

$$\begin{aligned} N^4 &= k^2(1-\gamma), \\ 2N^2 n_1^2 &= \big(\gamma(1+k) - 2k\big)(1-k), \\ 2N^2 n_3^2 &= \big(2k - \gamma(k+l)\big)(k-l), \\ n_1^2 &= |1-k|, \\ n_3^2 &= |k-l|. \end{aligned}$$

The last four of these give

$$N^2 = \left|\frac{\gamma(1+k) - 2k}{2}\right| = \left|\frac{2k - \gamma(k+l)}{2}\right|, \tag{6}$$

which leads to $\gamma = \frac{4k}{1+2k+l}$. Using (6) and $N^4 = k^2(1-\gamma)$, we arrive at $l = k^2$. Hence the ratios of the eigenvalues of A are equal and $\gamma = \frac{4k}{(1+k)^2} = \cos^2 \phi(1, k)$.

If n_1 or n_3 is the zero coordinate instead of n_2, the above argument shows that l is k^{-1} or $\sqrt{k}$, respectively, contradicting the assumption that $A_1 \leq A_2 \leq A_3$.

From above proof, when $1 < k$ we have $(n_1, 0, n_3) = (\pm\sqrt{k-1}, 0, \pm\sqrt{k^2-k}) \propto (\pm 1, 0, \pm\sqrt{k})$. So the four circles have normals $(\pm 1, 0, \pm\sqrt{k})$. Note that the level circle with normal $(1, 0, \sqrt{k})$ passes through $(\sqrt{k}, 1, 0)$ and $(0, \sqrt{k}, 1)$, corresponding to the two points of tangency of the parabola with the standard simplex. It is amusing to observe that as $k \to 1$, this circle goes to the circle on S^2 with normal $(1, 0, 1)$ and angular radius $\pi/3$.

4.3 Parabola

We now derive a formula for the conic $\mathcal{P} \cap \mathcal{M}(A, \gamma)$. For convenience, let $\vec{1} := (1, 1, 1)$, $\vec{A} := (A_1, A_2, A_3)$ and $\vec{A}^2 := (A_1^2, A_2^2, A_3^2)$. Define vectors

$$l = \vec{A} - \tfrac{1}{3}(\vec{A} \cdot \vec{1})\vec{1}, \qquad n = \vec{1} \times \vec{l} = \vec{1} \times \vec{A}.$$

Then $\{\vec{1}, l, n\}$ is an orthogonal basis of $\mathbb{R}^3$, and $\{l, n\}$ is an orthogonal basis of the space of vectors parallel to $\mathcal{P}$. Therefore $l \cdot s$ and $n \cdot s$ are orthogonal coordinates of $\mathcal{P}$.

Proposition 4.2 *The equation $s^T M_A(\gamma) s = 0$ of $\mathcal{M}(A, \gamma) \subset \mathcal{P}$ is $(l \cdot s - L)^2 = Mn \cdot s - N$, where*

$$L = \frac{\gamma}{2} \frac{l \cdot \vec{A}^2}{l \cdot l} - \frac{\vec{A} \cdot \vec{1}}{3}, \qquad M = \gamma \frac{n \cdot \vec{A}^2}{n \cdot n},$$

$$N = \frac{\gamma}{3} \frac{l \cdot \vec{A}^2}{l \cdot l} (\vec{A} \cdot \vec{1}) - \gamma \frac{\vec{1} \cdot \vec{A}^2}{\vec{1} \cdot \vec{1}} - \frac{\gamma^2}{4} \Big(\frac{l \cdot \vec{A}^2}{l \cdot l} \Big)^2.$$

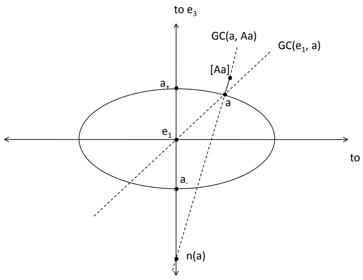
Proof Since the plane $\mathcal{P}$ is defined by $\vec{1} \cdot s = 1$, the equation $s^T M_A(\gamma) s = 0$ of $\mathcal{M}(A, \gamma) \subset \mathcal{P}$ may be written as $(\vec{A} \cdot s)^2 = \gamma(\vec{A}^2 \cdot s)$. This equation defines a parabola because for A_1, A_2 and A_3 distinct, $\vec{A} \cdot s$ and $\vec{A}^2 \cdot s$ are independent linear coordinates on $\mathcal{P}$. Since only the square of $\vec{A} \cdot s$ appears in the above equation the axis of the parabola is perpendicular to $\vec{A}$, and since it lies on $\mathcal{P}$ it is parallel to $\vec{n}$. Note that the axial direction is independent of γ. To obtain the desired equation, write $\vec{A}$ and $\vec{A}^2$ in terms of $\vec{1}, l$, and n, use $\vec{1} \cdot s = 1$, and complete the square appropriately.

Corollary 4.3 *$\mathcal{P} \cap \mathcal{M}(A, \gamma)$ is a parabola with vertex coordinates $l \cdot s = L$ and $n \cdot s = \frac{M}{N}$, and axis $l \cdot s = L$.*

5 Interactions Between the Level Curves

By Proposition 3.3, Problem 1 would be solved if the expression for the trace of $\phi_{S_\perp(a,b)}$ computed in Proposition 3.2 could be maximized. However, we were unable to carry out this maximization. Here we obtain a preliminary result using a geometric approach.

Throughout this section, suppose that A and B are 3×3 increasing diagonal PDS matrices with distinct eigenvalues. We work exclusively on S^2. We consider the level curves of the angle functions of A and B for fixed small angles α and β, respectively.



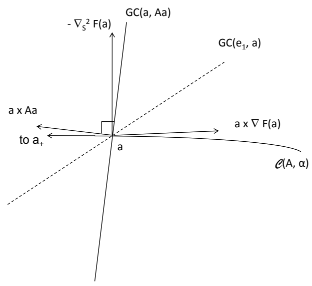


Figure 7: Polar projection of region of sphere near e_1.

Recall that level curves change qualitatively when the angles are either $\phi(A_1, A_2)$, $\phi(A_2, A_3)$ or $\phi(A_1, A_3) = \phi_A$. Choose α smaller than both $\phi(A_1, A_2)$ and $\phi(A_2, A_3)$. Similarly, choose β smaller than both $\phi(B_1, B_2)$ and $\phi(B_2, B_3)$. For such α, we saw that $\mathcal{L}(A, \cos^2 \alpha)$ consists of six disjoint simple closed curves around $\pm e_1, \pm e_2$ and $\pm e_3$.

Consider the portion of $\mathcal{L}(A, \cos^2 \alpha)$ around e_1. Let us call this component of the level curve $\mathcal{C}(A, \alpha)$. Note that $\mathcal{C}(A, \alpha)$ is symmetric about both $\mathrm{GC}(e_1, e_2)$ and $\mathrm{GC}(e_1, e_3)$. Let $a_\pm$ be the points on the intersection of $\mathrm{GC}(e_1, e_3)$ and $\mathcal{C}(A, \alpha)$ closest to e_1 in the northern and southern hemisphere, respectively (see Figure 7, left). Define $b_\pm$ similarly.

Theorem 5.1 *The minimum of $\phi_{S_{max}(a,b_-)}$ as a ranges over $\mathcal{C}(A, \alpha)$ is $\angle(a_+, b_-)$ and occurs at a_+.*

Recall that $\angle(a, Aa) = \alpha$ for all a on $\mathcal{C}(A, \alpha)$. We use again the notation $[x] = x/|x|$ in figures. If a lies on either $\mathrm{GC}(e_1, e_3)$ or $\mathrm{GC}(e_1, e_2)$, then $[Aa]$ also does so. Suppose that a does not lie on either $\mathrm{GC}(e_1, e_3)$ or $\mathrm{GC}(e_1, e_2)$. Without loss of generality, we may assume a to be a point on the first octant of S^2. A polar projection of the region of S^2 near e_1 is shown in Figure 7, left. The dotted line segments represent $\mathrm{GC}(e_1, a)$ and $\mathrm{GC}(a, Aa)$, as labeled.

Any two distinct great circles intersect each other at exactly two points. Consider the points of intersection of $\mathrm{GC}(e_1, e_3)$ and $\mathrm{GC}(a, Aa)$. Since $\mathrm{GC}(e_1, e_3)$ has axis e_2 and $\mathrm{GC}(a, Aa)$ has axis $a \times Aa$, these points of intersection are along $\pm e_2 \times (a \times Aa) = \pm(A_2 - A)a$. Let $n(a) = [(A_2 - A)a]$.

Lemma 5.2 *$n(a)$ is always between e_1 and $-e_3$ on $GC(e_1, e_3)$.*

Proof By definition, $n(a)$ is along $\begin{pmatrix} (A_2-A_1)a_1 \\ 0 \\ (A_2-A_3)a_3 \end{pmatrix}$. Since A is increasing diagonal with distinct eigenvalues, the third component of $n(a)$ is negative, and the first component is positive.

By the above lemma, $\mathrm{GC}(a, Aa)$ is always "steeper" than $\mathrm{GC}(e_1, a)$, as shown in Figure 7, left.

Next, recall from the previous section that $F(u) = \cos^2 \angle(u, Au)$, and $F = \cos^2 \alpha$ defines the level curve $\mathcal{L}(A, \cos^2 \alpha)$. Let $\nabla_{S^2} F$ be the projection of this vector to the tangent plane of the sphere at a. It is perpendicular to $\mathcal{C}(A, \alpha)$ at a and points inward from the oval, in the direction of decreasing $\angle(a, Aa)$. A polar projection near a is shown in Figure 7, right. Note that $a \times \nabla F$ is tangent to $\mathcal{C}(A, \alpha)$ at a and points away from a_+. The next proposition justifies the fact that in the drawing, $-\nabla_{S^2} F(a)$ is to the left of $\mathrm{GC}(a, Aa)$.

Proposition 5.3 *The great circle through a that is perpendicular to $\mathcal{C}(A, \alpha)$ passes through $GC(e_1, e_3)$ strictly between $n(a)$ and $-e_3$.*

Proof We first prove the following identity: $-\nabla F \cdot (a \times Aa) > 0$. Since at the point a, $\nabla F = 4\frac{(a \cdot Aa)Aa}{|Aa|^2} - 2\frac{(a \cdot Aa)^2 A^2 a}{|Aa|^4}$, we have

$$a \times \nabla F = 4\frac{a \cdot Aa}{|Aa|^2}(a \times Aa) - 2\frac{(a \cdot Aa)^2}{|Aa|^4}(a \times A^2 a).$$

Since $Aa \cdot (a \times Aa) = 0$, we have

$$Aa \cdot (a \times \nabla F) = -2\frac{(a \cdot Aa)^2}{|Aa|^4}\big\{Aa \cdot (a \times A^2 a)\big\}.$$

Now $A^2 a = (A_1^2 a_1, A_2^2 a_2, A_3^2 a_3)$, so $a \times A^2 a = \begin{pmatrix} a_2 a_3 (A_3^2 - A_2^2) \\ a_1 a_3 (A_1^2 - A_3^2) \\ a_1 a_2 (A_2^2 - A_1^2) \end{pmatrix}$. Thus

$$\begin{aligned} Aa \cdot (a \times A^2 a) &= a_1 a_2 a_3 \big\{A_1(A_3^2 - A_2^2) + A_2(A_1^2 - A_3^2) + A_3(A_2^2 - A_1^2)\big\}. \\ &= a_1 a_2 a_3 (A_3 - A_2)(A_2 - A_1)(A_1 - A_3). \end{aligned}$$

Since $A_1 < A_2 < A_3$, this quantity is negative, so $Aa \cdot (a \times \nabla F)$ and hence $-\nabla F \cdot (a \times Aa)$ are positive, proving the identity.

Using $a \cdot (a \times Aa) = 0$ and $a \cdot a = 1$, we obtain

$$\{(a \times Aa) \times a\} \cdot (a \times \nabla F) = -\nabla F \cdot (a \times Aa).$$

Since $a \times Aa$ is perpendicular to $\mathrm{GC}(a, Aa)$ at a, $(a \times Aa) \times a$ is parallel to $\mathrm{GC}(a, Aa)$ at a. Thus the angle measured counterclockwise from $(a \times Aa) \times a$ to $\nabla_{S^2} F$ is positive. This together with Lemma 5.2 proves the Proposition.

Hence for any a on $\mathcal{C}(A, \alpha)$, $\mathrm{GC}(a, Aa)$ is "between" $\mathrm{GC}(e_1, a)$ and the normal on the sphere to $\mathcal{C}(A, \alpha)$. Next we prove following lemma:

Lemma 5.4 *As a moves towards a_+ on $\mathcal{C}(A, \alpha)$, $n(a)$ moves monotonically towards $-e_3$.*

Proof Since $n(a)$ is along $\begin{pmatrix} (A_2-A_1)a_1 \\ 0 \\ (A_2-A_3)a_3 \end{pmatrix}$, $\left|\frac{n_3}{n_1}\right|$ is proportional to $\left|\frac{(A_2-A_3)a_3}{(A_2-A_1)a_1}\right|$. As A is fixed, the ratio $\left|\frac{n_3}{n_1}\right|$ is proportional to $\frac{a_3}{a_1}$. Hence as a moves towards a_+ on $\mathcal{C}(A, \alpha)$, $n(a)$ moves monotonically towards $-e_3$ if and only if $\frac{a_3}{a_1}$ is strictly increasing, or equivalently, $\frac{s_3}{s_1} = \frac{a_3^2}{a_1^2}$ is strictly increasing.

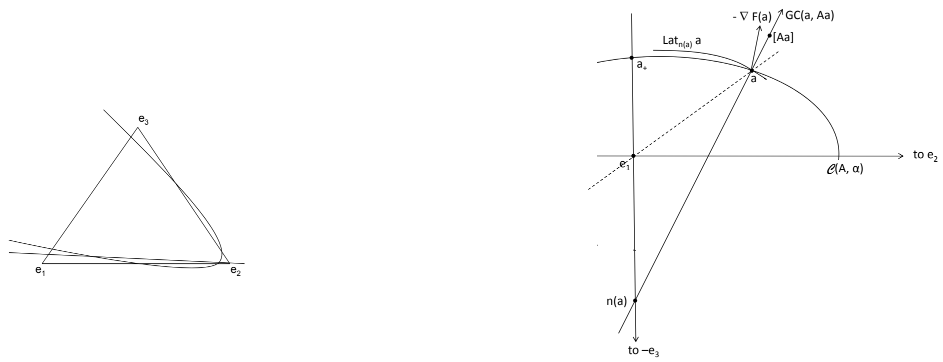


Figure 8: left: Schematic diagram, right: Polar projection at a.

Suppose that $\frac{s_3}{s_1}$ is not strictly increasing. Then for some point a on $\mathcal{C}(A, \alpha)$, there exists a second point $\tilde{a}$ on $\mathcal{C}(A, \alpha)$, between a and a_+, such that $\frac{\tilde{s_3}}{\tilde{s_1}} = \frac{s_3}{s_1}$. Note that $\frac{x_3}{x_1} = \frac{a_3}{a_1}$ defines $\mathrm{GC}(a, e_2)$ on the sphere, and $\frac{s_3}{s_1} = \frac{a_3^2}{a_1^2}$ defines the line segment passing through s and e_2 in the triangle $\mathcal{T}$. Since $\alpha < \min\{\phi(A_1, A_2), \phi(A_2, A_3)\}$, this line segment intersects the parabola corresponding to $\mathcal{C}(A, \alpha)$ at s and also at a point corresponding to a point on the oval on S^2 around e_2 (see Figure 8, left). Since $\frac{\tilde{s_3}}{\tilde{s_1}} = \frac{s_3}{s_1}$, the line segment also passes through $\tilde{s}$. This is a contradiction as a line segment intersects a parabola at most twice.

Define $\mathrm{Lat}_b(a)$ to be the latitude line through a with b as a pole, the path of a through all rotations about b. As a corollary to Proposition 5.3, we prove following lemma:

Lemma 5.5 *Consider the acute arc of* $\mathrm{Lat}_{n(a)}(a)$ *from* a *to* $GC(e_1, e_3)$. *Near* a*, this arc is on the outside of* $\mathcal{C}(A, \alpha)$ *(see Figure 8, right).*

Proof By Proposition 5.3, $\mathrm{GC}(a, Aa)$ is "between" $\mathrm{GC}(e_1, a)$ and the normal on the sphere to $\mathcal{C}(A, \alpha)$. Since $\mathrm{GC}(a, Aa)$ is perpendicular to $\mathrm{Lat}_{n(a)}(a)$, it must cross $\mathcal{C}(A, \alpha)$ at a from lower right to upper left as shown in Figure 8, right.

Lemma 5.6 $\mathrm{Lat}_{n(a)}(a)$ *never crosses* $\mathcal{C}(A, \alpha)$ *between* a_+ *and* a.

Proof We prove this by contradiction. Suppose that $\mathrm{Lat}_{n(a)}(a)$ crosses $\mathcal{C}(A, \alpha)$ at a point $\tilde{a}$ between a_+ and a from above. By Lemma 5.4, $n(\tilde{a})$ is between $n(a)$ and $-e_3$. Let $\tilde{S}$ be the rotation around $n(a)$ such that $\tilde{a} = \tilde{S}a$. Then $[\tilde{S}Aa]$ lies on $\mathrm{GC}(\tilde{a}, n(a))$. Note that $\mathrm{GC}(\tilde{a}, n(a))$ is perpendicular to $\mathrm{Lat}_{n(a)}(a)$, and $\nabla_{S^2} F(\tilde{a})$ is perpendicular to the tangent to $\mathcal{C}(A, \alpha)$ at $\tilde{a}$. Thus the angle measured counterclockwise from the tangent to $\mathcal{C}(A, \alpha)$ at $\tilde{a}$ to $\mathrm{GC}(\tilde{a}, n(a))$ is obtuse (see Figure 9, left). This contradicts Proposition 5.3. Hence $\mathrm{Lat}_{n(a)}(a)$ never crosses $\mathcal{C}(A, \alpha)$ from above between a and a_+. Therefore it never crosses $\mathcal{C}(A, \alpha)$ between a and a_+ at all, as it starts out above at a by Lemma 5.5. Similarly, it can never be tangent to $\mathcal{C}(A, \alpha)$ between a and a_+.

Thus the situation is as shown below in Figure 9, right.

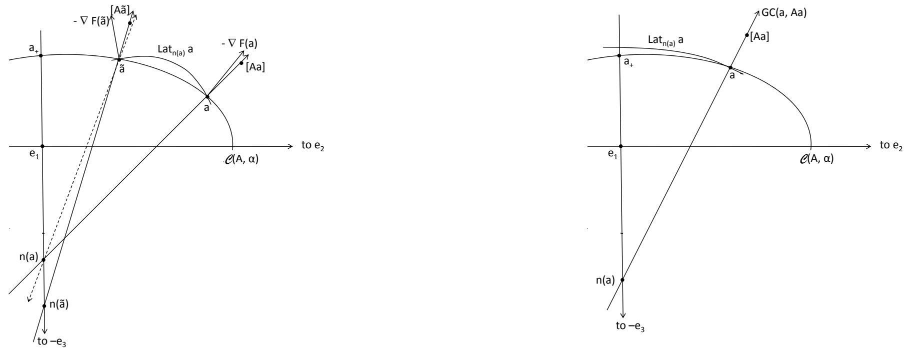


Figure 9: Polar projection at e_1.

We need two elementary results before proving Theorem 5.1.

Lemma 5.7 *Let* $S_1, S_2 \in \mathrm{SO}_3$ *have perpendicular axes of rotation. Then*

$$\cos^2\left(\frac{\phi_{S_1 S_2}}{2}\right) = \cos^2\left(\frac{\phi_{S_1}}{2}\right)\cos^2\left(\frac{\phi_{S_2}}{2}\right).$$

Proof Without loss of generality, we may assume $S_1 = R(e_3, \phi_{S_1})$ and $S_2 = R(e_2, -\phi_{S_2})$. Computation gives $\mathrm{Tr}(S_1S_2) = (1 + \cos\phi_{S_1})(1 + \cos\phi_{S_2}) - 1$. On the other hand $\mathrm{Tr}(S_1S_2) = 1 + 2\cos\phi_{S_1S_2}$. The result follows.

Corollary 5.8 *Let $S_1, S_2 \in \mathrm{SO}_3$ be two matrices with perpendicular axes of rotation. Then $\phi_{S_1S_2} \geq \max\{\phi_{S_1}, \phi_{S_2}\}$.*

Proof of Theorem 5.1. Consider the point b_- on the intersection of $\mathcal{C}(B, \beta)$ and $\mathrm{GC}(e_1, e_3)$ between e_1 and $-e_3$. Recall that $S_{\max}^{-1}(a, b_-)$ is the orthogonal matrix that moves a to b_- such that $[S_{\max}^{-1}(a, b_-)Aa]$ lies on $\mathrm{GC}(e_1, e_3)$, opposite $[Bb_-]$ from b_- (see Figure 10, left).

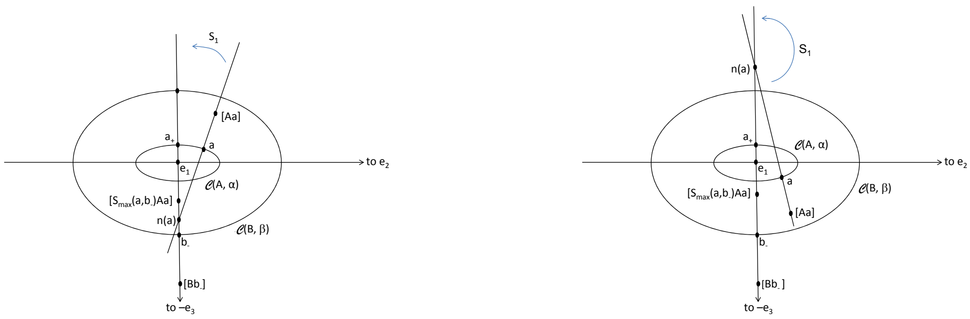


Figure 10: Polar projection of region of sphere near e_1.

Note that $\phi_{S_{\max}(a_+, b_-)} = \angle(a_+, b_-)$. By Corollary 2.2, $\angle(a_+, e_1) \leq \frac{\pi}{4} - \frac{\phi_A}{2}$ and $\angle(e_1, b_-) \leq \frac{\pi}{4} - \frac{\phi_B}{2}$, so both $\angle(a_+, e_1)$ and $\angle(e_1, b_-)$ are less than or equal to $\frac{\pi}{4}$. Therefore $\angle(a_+, b_-) \leq \frac{\pi}{2}$.

First suppose that a is a point on $\mathcal{C}(A, \alpha)$ in the first octant of S^2 (i.e., the first quadrant in Figure 10). We factor $S_{\max}^{-1}(a, b_-)$ as S_2S_1, where S_1 and S_2 are the following rotations: S_1 rotates $\mathrm{GC}(a, Aa)$ to $\mathrm{GC}(e_1, e_3)$ counterclockwise (hence by an acute angle) about $n(a)$, and S_2 moves $[S_1a]$ to b_- about e_2 so that $[S_2S_1Aa]$ is opposite $[Bb_-]$ on $\mathrm{GC}(e_1, e_3)$.

By Lemma 5.6, since $\mathrm{Lat}_{n(a)}(a)$ never crosses $\mathcal{C}(A, \alpha)$ between a and a_+, S_1a is above a_+ on $\mathrm{GC}(e_1, e_3)$ (see Figure 9, right). Thus the distance between S_1a and b_- is greater than the distance between a_+ and b_- along $\mathrm{GC}(e_1, e_3)$, so $\phi_{S_2} > \angle(a_+, b_-)$. By Corollary 5.8, $\phi_{S_{\max}(a,b_-)} \geq \max\{\phi_{S_1}, \phi_{S_2}\}$. Therefore $\phi_{S_{\max}(a,b_-)} > \angle(a_+, b_-)$.

Next, suppose a is on the part of $\mathcal{C}(A, \alpha)$ in the fourth quadrant (Figure 10, right). We factor $S_{\max}^{-1}(a, b_-)$ as S_2S_1, where S_1 and S_2 are the following rotations: S_1 rotates $\mathrm{GC}(a, Aa)$ to $\mathrm{GC}(e_1, e_3)$ counterclockwise (hence by an obtuse angle) about $n(a)$, and S_2 rotates $[S_1a]$ to b_- about e_2 so that $[S_2S_1Aa]$ is opposite $[Bb_-]$ on $\mathrm{GC}(e_1, e_3)$. Since $\phi_{S_1} > \frac{\pi}{2}$, by Corollary 5.8 $\phi_{S_{\max}(a,b_-)} > \frac{\pi}{2}$.

The proof is similar if a is on the second or the third quadrant. □

It is natural to expect that when A and B are both diagonal with increasing entries, the solution of Problem 1 is the same as for the matrices $\begin{pmatrix} A_1 & \\ & A_3 \end{pmatrix}$ and $\begin{pmatrix} B_1 & \\ & B_3 \end{pmatrix}$. We conclude with a conjectural generalization of Theorem 5.1 which would imply this.

Conjecture 5.9 *The minimum of* $\phi_{S_{\max}(a,b)}$ *on* $\mathcal{C}(A,\alpha)\times\mathcal{C}(B,\beta)$ *occurs at* (a_+, b_-) *and* (a_-, b_+).

With greatest regard, I would like to thank Dr. Charles Conley, my major advisor for all the patience, time and scholarly insight provided by him during this study. This work is the result of his guidance and constant encouragement.

References

[1] C. H. Conley, P. Pucci, and J. Serrin, Elliptic equations and products of positive definite matrices, *Mathematische Nachrichten,* **21**, (2005), 278.

[2] D. W. Nicholson, Eigenvalue bounds for $AB + BA$, with A, B positive definite matrices, *Linear Algebra and its Application*, **24**, (1979), 173-183.

[3] G. Strang, Eigenvalues of Jordan Products, *The American Mathematical Monthly,* **69**, (1962).

[4] N. Alikakos and P. W. Bates, Estimates for the Eigenvalues of the Jordan Product of Hermitian Matrices, *Linear Algebra and its Application,* **57**, (1984), 41-56.

[5] O. Taussky Todd, Research Problem 2, *Bull. Amer. Math. Soc.,* **66**, (1960), 275.

[6] R. A. Horn and C. A. Johnson, *Matrix Theory,* Cambridge University Press, (1985).

Received: October, 2010