2002

Determining If Two Ellipsoids Share the Same Volume

Salvatore Alfano

Meredith L. Greer
Bates College, mgreer@bates.edu

Recommended Citation

Alfano, S., Greer, M.L., 2002, Determining If Two Ellipsoids Share the Same Volume, Advances in the Astronautical Sciences, 109, 771-783.

(both cross-track and radial) was only, 0.0209 km., with the angle of the drag vector to the velocity vector being 0.29 degrees. For a 10% increase in drag, the increase in drag retardation was 0.415110 km., and the out-of-track angle was 0.27 degrees. This indicates that the along-track approximation is typically good to a fraction of a percent. The small out-of-track component is likely due to the fact that the drag is actually with respect to an atmosphere which co-rotates with the earth to a first approximation. This results in a small out-of-track drag retardation.

The clear conclusion is that, to the extent atmospheric drag is the dominant prediction error source, the errors in the along-track positions of two objects are correlated. This is the case now for many LEO objects, and will be the case increasingly in the future as orbital determination techniques improve. It means that, where the drag dominates, the error volume approaches a curve in space with the equation above, and independent variable $\delta\rho$, the atmospheric density variation. The actual error bound will have a cross dimension of the cross-track error, which is normally only weakly dependent on drag.

The previous analysis has taken an experimental physicist's approach, with little mention of the underlying dynamical equations. This has been done purposely to emphasize the physical traceability of the errors discussed without getting distracted by complex mathematics. There is an important conceptual difference between this approach and the commonly used method of using the covariance matrix of the solution to predict position errors. The covariance method uses mathematical analysis to relate errors from all causes in the orbit determination to possible uncertainty in the predicted position of the RSO solved for. This analysis, in contrast, uses knowledge of the physics of atmospheric drag retardation and its effect on both objects that could only be put formally into the methematical model by a simultaneous solution. Even in this case, the constancy of the B term is determined by additional longer term analysis. To summarize, the optimum prediction of the actual error volume in possible collisions will be obtained by use of this simple drag-related variance algorithm, together with accuracy estimation for the atmospheric drag model in use which can be supplemented with real time updates. This needs to be coupled with associated knowledge of the constancy of the B term for the two RSOs. Is also needs to be supplemented by estimates of the errors in position from all other sources including sensor accuracy.

AAS 01-357

DETERMINING IF TWO ELLIPSOIDS SHARE THE SAME VOLUME*

Salvatore Alfano† and Meredith L. Greer†

An analytical method is presented for determining if two ellipsoids share the same volume. The formulation involves adding an extra dimension to the solution space and examining eigenvalues that are associated with degenerate quadric surfaces. The eigenvalue behavior is characterized and then demonstrated with an example. The same method is also used to determine if two ellipsoids appear to share the same projected area based on an oberver's viewing angle. The following approach yields direct results without approximation, iteration, or any form of numerical search. It is computationally efficient in the sense that no dimensional distortions, coordinate rotations, transformations, or eigenvector computations are needed.

† The Aerospace Corporation.

Introduction

As the US Satellite Catalog transitions from General Perturbations to Special Perturbations, the positional accuracy of each space object will be readily available in the form of a covariance matrix. These covariances can be used to determine probability of collision, radio frequency interference, and/or incidental laser illumination. Because the probability calculations can be computationally burdensome, it is desirable to prescreen candidate objects based on user-defined thresholds. Specifically, each object can be represented by a covariance-based ellipsoid and then processed to determine if its uncertainty volume shares some space in common with another's. Ellipsoids (or their projections) that do not touch or overlap can be eliminated from further processing. This article presents a simple analytical method to perform such screening.

To date, all ellipsoidal prescreening methods involve numerical searches (Ref. 1). For computational efficiency, such prescreening is often reduced to spheres or "keep-out" boxes that have much larger volumes but allow for quick distance comparisons. The drawback to such screening is that these larger volumes cause many objects to become candidates for further (albeit unnecessary) processing. These methods result in increased downstream computational processing and /or increased operator workload to further assess potential satellite conjunctions.

The following method adds an extra dimension to the solution space. The subset of eigenvalues that are associated with intersecting degenerate quadric surfaces are then examined. The same method is also used to determine if two ellipsoids appear to share the same projected area based on viewing angle. The approach yields direct results without approximation, iteration, or any form of numerical search. It is computationally efficient in the sense that no dimensional distortions, coordinate rotations, transformations, or eigenvector computations are needed. This method expands the two-dimensional work of Hill (Ref. 2) in his formulation of degenerate conics. It also furthers his work by examining the associated eigenvalue behavior.

This approach is not limited to Satellite Catalog applications. For computer graphics users, such screening could be used to invoke a hidden line removal algorithm.

Ellipsoidal Formulation

Rogers and Adams (Ref. 3) give various representational forms for an ellipsoid. Algebraically, the representation is

$$Ax^2 + By^2 + Cz^2 + Dxy + Eyz + Fxz + Gx + Hy + Jz + K = 0 \qquad (1)$$

where A, B, C, D, E, F, G, H, J, and K are constants. In matrix form, the same ellipsoid can be written as

$$X \cdot S \cdot X^T = 0 \qquad (2)$$

where

$$X = [x\ y\ z\ 1] \qquad (3)$$

$$S = \frac{1}{2} \cdot \begin{pmatrix} 2A & D & F & G \\ D & 2B & E & H \\ F & E & 2C & J \\ G & H & J & 2K \end{pmatrix} . \qquad (4)$$

The translation of the ellipsoid's center from the origin to [X0, Y0, Z0] can be accomplished by the matrix

$$T = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -X0 & -Y0 & -Z0 & 1 \end{pmatrix} \qquad (5)$$

where

$$X \cdot T \cdot S \cdot T^T \cdot X^T = 0 \ . \qquad (6)$$

Similarly, all points contained within the ellipsoid satisfy the constraint

$$X \cdot T \cdot S \cdot T^T \cdot X^T \leq 0 \ . \qquad (7)$$

Given a 3x3 covariance matrix C centered about [X0, Y0, Z0], the quadric representation of the ellipsoid would then be

$$X \cdot T \cdot \begin{pmatrix} Ci_{11} & Ci_{12} & Ci_{13} & 0 \\ Ci_{21} & Ci_{22} & Ci_{23} & 0 \\ Ci_{31} & Ci_{32} & Ci_{33} & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \cdot T^T \cdot X^T = 0 \qquad (8)$$

where Ci are the elements of the inverted covariance matrix.

Ellipsoidal Solution

For simplicity, assume a primary object is centered at the origin. An ellipsoid that corresponds to its positional covariance can be computed from the above, resulting in the equation

$$X \cdot A \cdot X^T = 0 \quad . \qquad (9)$$

In the same manner, a secondary object (center not co-located) and its ellipsoid can be appropriately translated relative to the primary object such that

$$X \cdot B \cdot X^T = 0 \quad . \qquad (10)$$

If any X exists such that it satisfies Eqs. (9) and (10), then the primary and secondary ellipsoids intersect at that point. If some value of X satisfies the constraint for both objects as represented by Eq. (7), then that point lies inside both ellipsoids.

Eq. (9) can be multiplied by a scalar constant λ with no loss in generality; the constant is then brought inside the equation.

$$\lambda \cdot X \cdot A \cdot X^T = 0 \qquad (11)$$

$$X \cdot (\lambda \cdot A) \cdot X^T = 0 \qquad (12)$$

Assuming a subset of X satisfies both Eqs. (9) and (10), then it must also satisfy any linear combination of the two. Such a combination is shown by differencing Eqs. (10) and (12) to produce

$$X \cdot (\lambda \cdot A - B) \cdot X^T = 0 \quad . \qquad (13)$$

As explained by Hill (Ref. 2), λ is chosen so that the parenthetical term is degenerate; this occurs when its determinant is zero. Because A is the

characteristic matrix of an ellipsoid, it is invertible and can be used to alter Eq. (13) to produce

$$X \cdot A \cdot \left(\lambda I - A^{-1} \cdot B\right) \cdot X^T = 0 \quad . \qquad (14)$$

This representation is more readily recognized as an eigenvalue formulation and also lends itself well to many mathematical software packages.

Substituting selected eigenvalues into Eq. (14) will produce characteristic matrices that represent degenerate quadric surfaces. If the X subset assumption holds regarding overlapping objects, then these surfaces must also pass through the points shared by the primary and secondary ellipsoids. It can be deduced that if the ellipsoids just touch (i.e. share a single point in common) then that solution vector must also be an eigenvector of $A^{-1} \cdot B$ (Ref 4). The converse is not true as not all eigenvectors of $A^{-1} \cdot B$ will satisfy the ellipsoidal constraints of Eqs. (9) and (10). Eigenvectors with a zero in their last component are considered inadmissible because this formulation has been framed in a four-dimensional space with the last dimension fixed as shown in Eq. (3). An admissible eigenvector can be tested by simply scaling it to produce a one in the last component and then determining if it meets the ellipsoidal conditions as represented by matrices A and B.

When the primary and secondary ellipsoids overlap, then a family of solutions describes the intersection. For such cases, two of the eigenvalues become complex. This is demonstrated in Appendix A and proven in (Ref 4).

Observed Eigenvalue Behavior

To gain an understanding of the eigenvalues when the ellipsoids don't just touch, the locus of values was plotted for various cases by altering size, shape, orientation, and location. Figure 1 is representative of all cases tested. In each set of cases, the two ellipsoids were initially defined to be completely outside each other. There were always two negative, real eigenvalues that produced admissible eigenvectors. The vectors did not satisfy Eqs. (9) and (10) and no point was shared in common between the ellipsoids.

The primary ellipsoid was continually scaled up until it just touched the secondary. This meant that only a single, unique point satisfied Eqs. (9) and (10). The two eigenvalues moved towards each other until they met (repeated). At this point the admissible eigenvectors gave the solution to where the ellipsoids touch.

The scaling then continued so that both ellipsoids shared some volume in common. The two admissible eigenvalues became complex conjugates. The real portion of the eigenvectors satisfied the inequality for both ellipsoids as defined in Eq. (7). The location indicated by these vectors was always shown to be inside both ellipsoids; therefore they intersected.

As the primary ellipsoid continued to grow, it eventually touched the far side of the secondary. The two admissible eigenvalues again became real and repeated, but were positive instead of negative. Again, those eigenvectors defined the exact point where the ellipsoids touched.

Scaling beyond this point always gave two positive real admissible eigenvalues that moved away from each other. In all cases tested it meant that some portion of the primary surface had entered and exited the secondary ellipsoid. It did not mean that the primary had completely engulfed the secondary. This example can best be visualized as a broom stick going completely through a rugby ball. A simplified mathematical explanation for eigenvalue behavior is presented in Appendix A. The complete, n-dimensional, mathematical proof is the subject of another paper (Ref. 4).

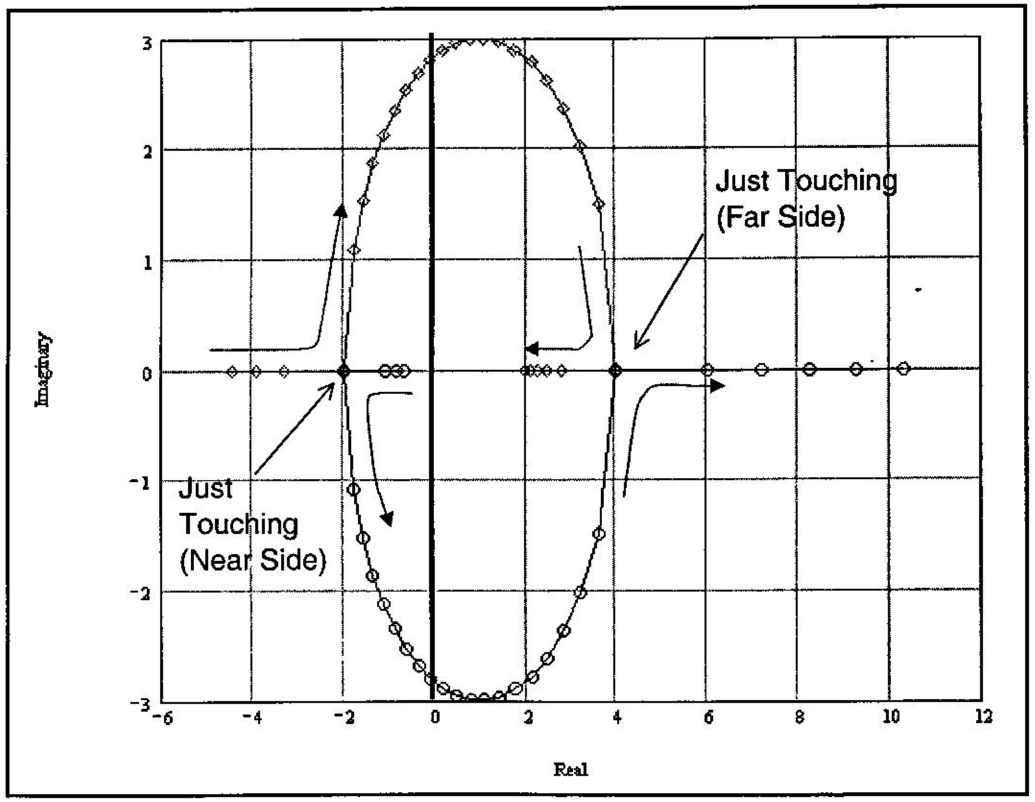

Figure 1. Representative locus of admissible eigenvalues

To summarize the results, the admissible eigenvalues of $A^{-1} \cdot B$ were examined. If two were negative real and different, then they shared no volume in

common. If two were negative real and identical, then they shared a single point in space (just touching on the secondary ellipsoid's side nearest the origin). If two were complex conjugates, the surfaces intersected. If two were positive real and identical, then they shared volume and a single point in space (just touching on the side farthest from the origin). If all were positive real, then one ellipsoid completely penetrated the other without necessarily engulfing it (as demonstrated in the following figure).

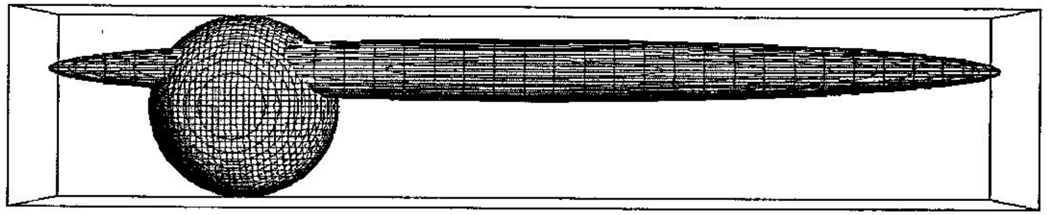

Figure 2. Complete penetration of one ellipsoid by another

Simple Ellipsoidal Example

This example involves a primary ellipsoid that is 4 units long on the x axis and 2 units long on the y and z axes. The secondary is 6 units long on the x axis, 4 on the y and 8 on the z with its center at [7, 0, 0]. The primary should just touch the secondary on the near side when scaled by 2 and just touch the far side when scaled by 5. The touching will occur on the x axis.

The initial A and B matrices are

$$A = \begin{pmatrix} 0.25 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{15}$$

$$B = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -7 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0.111 & 0 & 0 & 0 \\ 0 & 0.25 & 0 & 0 \\ 0 & 0 & 0.063 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & -7 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{16}$$

$$B = \begin{pmatrix} 0.111 & 0 & 0 & -0.778 \\ 0 & 0.25 & 0 & 0 \\ 0 & 0 & 0.063 & 0 \\ -0.778 & 0 & 0 & 4.444 \end{pmatrix} \tag{17}$$

Scaling the primary ellipsoid by a factor of n is done by simply multiplying the last element of A by n^2.

Table 1 shows the history of the eigenvalues and their interpretations.

Scale	1	2	3	4	5	6
Eig 1	-0.114	-0.333	-0.025+0.221i	.083+0.114i	0.133	0.276
Eig 2	-3.886	-0.333	-0.025-0.221i	.083-0.114i	0.133	0.045
Vector	N/A	$\begin{pmatrix}4\\0\\0\\1\end{pmatrix}$	$\begin{pmatrix}5.429-2.556i\\0\\0\\1\end{pmatrix}$	$\begin{pmatrix}7.429-2.969i\\0\\0\\1\end{pmatrix}$	$\begin{pmatrix}10\\0\\0\\1\end{pmatrix}$	N/A
Notes	Outside	Touch	Overlap	Overlap	Touch	Past

Table 1: Effects of scaling on eigenvalues and eigenvectors.

Coordinate Reduction through Projection

Although two ellipsoids may not share the same space, when viewed from certain angles one may appear to cover or overlap the other. Analysis of such circumstances is necessary to prevent accidental laser illumination if a secondary object is in or near the line of sight of the primary. Equally important is determining the possibility of radio frequency interference on a secondary object. For computer graphics users, such analysis would indicate when to invoke a hidden line removal algorithm. Coordinate rotations are accomplished through the following matrix representation

$$X \cdot R \cdot S \cdot R^T \cdot X^T = 0 \quad , \tag{18}$$

where rotation about the x axis of angle α produces

$$Rx = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos(\alpha) & \sin(\alpha) & 0 \\ 0 & -\sin(\alpha) & \cos(\alpha) & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} , \tag{19}$$

rotation about the y axis of angle β yields

$$Ry = \begin{pmatrix} \cos(\beta) & 0 & -\sin(\beta) & 0 \\ 0 & 1 & 0 & 0 \\ \sin(\beta) & 0 & \cos(\beta) & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} , \tag{20}$$

and rotation about the z axis of angle θ is

$$Rz = \begin{pmatrix} \cos(\theta) & \sin(\theta) & 0 & 0 \\ -\sin(\theta) & \cos(\theta) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} . \tag{21}$$

The individual matrices can be multiplied to produce an overall rotation matrix R. The reader is cautioned to pay close attention to the signs of the sine terms; this is necessary for a positive right-hand rule convention. Also, the order of multiplication is important to assure the desired overall coordinate rotation.

Coordinate reduction is done by means of a simple orthographic projection in the rotated space to eliminate one component. The choice of coordinate for reduction is a matter of personal preference. The new z component was chosen for this work, resulting in

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{22}$$

$$X \cdot \left(P \cdot R \cdot S \cdot R^T \cdot P^T\right) \cdot X^T = 0 \quad . \tag{23}$$

When the projection is completed, the expression in parentheses becomes singular. To proceed, it is necessary to reduce the dimension of the state vector and associated formulation as will be explained in the next section.

It is still necessary to translate the resultant based on the new coordinate frame. To do so, a new translation vector is computed and inserted into the translation matrix

$$[X1 \;\; Y1 \;\; Z1 \;\; 1] = [X0 \;\; Y0 \;\; Z0 \;\; 1] \cdot R \tag{24}$$

$$\text{Tnew} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -X1 & -Y1 & -Z1 & 1 \end{pmatrix} \quad . \tag{25}$$

Combining all terms in the correct order produces

$$X \cdot \text{Tnew} \cdot P \cdot R \cdot S \cdot R^T \cdot P^T \cdot \text{Tnew}^T \cdot X^T = 0 \quad . \tag{26}$$

Elliptical Formulation and Solution

As one would expect, determining if two ellipses share the same area is identical to the ellipsoidal formulation reduced by one dimension. In matrix form, the new z component resulting from coordinate rotation is eliminated and the equations are reduced by one dimension such that

$$X = [x\ y\ 1] \quad . \tag{27}$$

An ellipsoid described by the rotated 4x4 A matrix is projected into the new x-y plane by removing the third row and column to produce the 3x3 AP matrix. The relationship

$$X \cdot AP \cdot X^T = 0 \tag{28}$$

now describes the primary object's projected ellipse in the new, dimensionally-reduced frame. The same projection and reduction is done for the secondary object to determine the BP matrix

$$X \cdot BP \cdot X^T = 0 \quad . \tag{29}$$

If any X exists such that it satisfies Eqs. (28) and (29), then the primary and secondary projections intersect at that point. If some value of X satisfies the constraint for both projections as represented by Eq. (7), then that point lies inside both ellipses.

The evaluation is identical to the ellipsoidal one, observing the admissible eigenvalue behavior of $AP^{-1} \cdot BP$ to determine if the ellipses shared the same space. If two are negative real and different, then the ellipses share no area in common. If two are negative real and identical, then they just touch on the secondary's side nearest the origin. If two are complex conjugates, the ellipses intersect at two points. If two are positive real and identical, then they share area

and just touch on the far side. If all are positive real, then one penetrates or engulfs the other.

Conclusions

A simple analytical method has been developed to determine if two ellipsoids share the same volume. This method can be used to alert operators of existing or impending conjunctions. The formulation involves adding an extra dimension to the solution space and examining the admissible eigenvalues. The admissible eigenvalues are examined to determine if any volume is shared. If volume is shared, a subset of the eigenvalues define degenerate quadric surfaces that pass through the points of intersection. The same method is used to determine if two ellipsoids appear to share the same projected area based on viewing angle. This approach yields direct results without approximation, iteration, or any form of search.

References

1. D. L. Oltrogge and R. G. Gist,., "Collision Vision: Covariance Modeling and Intersection Detection for Spacecraft Situational Awareness," AIAA/AAS Astrodynamics Specialist Conference, 16-19 August 1999, AIAA 99-351, pp. 5-7
2. K. Hill, "Matrix-Based Ellipse Geometry," *Graphics Gems V*, Academic Press, 1995, pp. 72-77
3. D. F. Rogers, and J. A. Adams, *Mathematical Elements for Computer Graphics*, 2nd ed., McGraw-Hill, 1990, pp. 400-408.
4. K. Chan, "A Simple Mathematical Approach for Determining Intersection of Quadratic Surfaces," AIAA/AAS Astrodynamics Specialist Conference, July 30 - August 2, 2001, AAS 01-358.

APPENDIX A

The mathematical underpinnings for the assertions of eigenvalue behavior in two and three dimensions are proven here for a single dimension; the n-dimensional proof is found in Ref. 4. All objects can be scaled and rotated so that the primary is centered at the origin with unit dimensions. The primary ellipsoid becomes a sphere, the primary ellipse a circle. By selecting the proper viewing geometry, two ellipsoids that do not touch can be projected to two ellipses that do not touch; these ellipses can then be projected to two lines that do not touch. This process reduces the problem to a single dimension.

For a single dimension, the primary object is a line ranging from –1 to +1 with its "surface" represented by the end points. The secondary is also a line ranging from (x0-a) to (x0+a). Scaling can be accomplished so that the only case needing consideration is when x0>0 and a>0. Algebraically these endpoints can be expressed as

$$x^2 = 1 \qquad \text{(A1)}$$

$$a^{-2}\cdot(x - x0)^2 = 1 \quad . \qquad \text{(A2)}$$

In matrix form these become

$$(x \ \ 1)\cdot\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\cdot\begin{pmatrix} x \\ 1 \end{pmatrix} = 0 \qquad \text{(A3)}$$

$$(x \ \ 1)\cdot\begin{pmatrix} 1 & 0 \\ -x0 & 1 \end{pmatrix}\cdot\begin{pmatrix} a^{-2} & 0 \\ 0 & -1 \end{pmatrix}\cdot\begin{pmatrix} 1 & -x0 \\ 0 & 1 \end{pmatrix}\cdot\begin{pmatrix} x \\ 1 \end{pmatrix} = 0 \quad . \qquad \text{(A4)}$$

The eigenvalues of (λA-B) are solved such that

$$A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \qquad \text{(A5)}$$

$$B = \begin{pmatrix} \dfrac{1}{a^2} & \dfrac{-x0}{a^2} \\ \dfrac{-x0}{a^2} & \dfrac{x0^2}{a^2} - 1 \end{pmatrix} \qquad \text{(A6)}$$

$$\lambda = \frac{-x0^2 + a^2 + 1 \pm \sqrt{(a+1-x0)\cdot(a+1+x0)\cdot(a-1-x0)\cdot(a-1+x0)}}{2\cdot a^2} \quad . \qquad \text{(A7)}$$

The following figure helps in visualizing all possible values, both real and complex, of the solution.

Figure 3: One dimensional analysis (a>0, x0>0)

Figure 3 shows that when [(x0 - a) > 1] the lines do not touch. Placing this constraint into Eq. (A7) will always produce negative, real, distinct eigenvalues.

Increasing the value [a] and/or decreasing the value [x0] such that [(x0 - a) = 1] allows the lines to just touch on the positive (near) side. The eigenvalues repeat with a value of [–1/a].

Continuing to increase [a] or decrease [x0] such that [-1 < (x0 - a) < 1] and [(x0 + a) > 1] causes the lines to overlap, but not completely. The eigenvalues will always be complex conjugates under these conditions.

Should [(x0 - a) = -1] and [(x0 + a) > 1] then the lines overlap and just touch on the negative (far) side. The eigenvalues repeat with a value of [+1/a].

In the event that [-1 < (x0 - a) < 1] while [(x0 + a) <= 1] then the secondary line is completely inside the primary and the eigenvalues are positive, real, and distinct.

For the final case [(x0 - a) < -1] the primary line is completely inside the secondary and the eigenvalues are again positive, real, and distinct.