Geometry of Continuous Adjoint Newton's Method for Bivariate Quadratics

Jonathan D. Hauenstein, Caroline Hills, Hoon Hong, and Francisco Ponce-Carrion

December 14, 2024

Abstract

Newton's method is a classical iterative approach for computing solutions to nonlinear equations. To overcome some of its drawbacks, one often considers a continuous adjoint form of Newton's method. This paper investigates the geometric structure of the trajectories produced by the continuous adjoint Newton's method for bivariate quadratics, a system of two quadratic polynomials in two variables, via eigenanalysis at its equilibrium points.

1 Introduction

In this section, we revisit the *continuous adjoint Newton's method* for solving system of equations. Specifically, we investigate the *geometric structure of its trajectories via eigenanalysis* at its equilibrium points. Below, we describe these two aforementioned italicized phrases.

Continuous adjoint Newton's method: We will elaborate on this method by recounting where it came from and why. See [5] for the history of Newton's method.

First, recall the well-known *Newton's method* for solving a well-constrained multivariate system $f(x) = 0$ begins with an initial guess x_0 that is iteratively updated by the expression

$$x_{j+1} = x_j - f'(x_j)^{-1} \cdot f(x_j)$$

where $f'(x)$ stands for the Jacobian matrix of $f(x)$. One downside of Newton's method is that the set of initial values that lead to a particular solution is usually fractal. This creates both chaos and beautiful pictures, e.g., see [3]. To overcome this difficulty, one can make infinitesimal updates, resulting in a *continuous Newton's method*, e.g., see [1], namely

$$\frac{dx}{dt} = -f'(x)^{-1} f(x).$$

The continuous Newton's method requires f' to be nonsingular and thus invertible, which is not always the case. In order to overcome this difficulty [1], we multiply both sides by $|f'|$, the determinant of f', and reparameterize t appropriately to obtain

$$\frac{dx}{dt} = g(x) \qquad \text{where} \qquad g = |f'|\, f'^{-1} f. \tag{1}$$

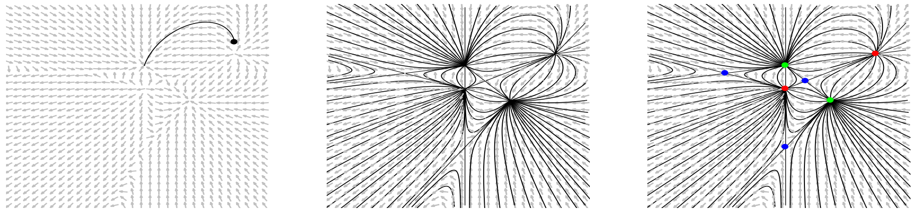
For simplicity, we denote $f = f(x)$ and so on. Note that $|f'|\, f'^{-1}$ is the adjoint of f'. Hence, to distinguish various continuous Newton's methods, (1) is called the *continuous adjoint Newton's method*.

Geometric structure of trajectories via eigenanalysis: To elaborate this phrase's meaning, we use the following example. It also serves as the system to demonstrate the theoretical results of this work.

Illustrative Example. Consider the bivariate quadratic system

$$f = \begin{bmatrix} -11x_1^2 + 8x_1x_2 - 2x_2^2 + 63x_1 - 112 \\ -8x_1^2 + 5x_1x_2 - 8x_2^2 + 54x_1 + 54x_2 - 196 \end{bmatrix}. \tag{2}$$

The following three plots are obtained from the continuous adjoint Newton's method (1):



(a) (b) (c)

Consider the plots one by one.

(a) The background provides the field plot (dim arrows) of g which show the direction of g at each grid point $x = (x_1, x_2)$. The foreground plots a trajectory (black curve) of $\frac{dx}{dt} = g(x)$ starting from an initial point (black point).

(b) Many trajectories (black curves) are plotted starting from different initial points (not shown). The structure of these trajectories is the main subject of this work.

(c) Seven points (in red, green, and blue) are highlighted. If one of these points is chosen as an initial point, the corresponding trajectory is stationary since $\frac{dx}{dt} = g = 0$. Hence, those seven points are the equilibria and the other trajectories not starting at an equilibria need to be investigated.

- All trajectories near a red equilibrium appear to go away from the equilibrium. Such an equilibrium is called a *source*.
- All trajectories near a green equilibrium appear to go toward the equilibrium. Such an equilibrium is called a *sink*.
- Almost all trajectories near a blue equilibrium appear to go toward the equilibrium and then away from it. Such an equilibrium is called a *saddle*.

In particular, the geometric structure of the trajectories can be obtained by analyzing the equilibria and determining their types: source, sink, and saddle. The Hartman-Grobman Theorem, e.g., see [2], posits that the type of each equilibrium can be determined by eigenanalysis of the Jacobian matrix g' at the equilibrium:

- if all the eigenvalues are positive then the equilibrium is a source,
- if all the eigenvalues are negative then the equilibrium is a sink, and
- if some eigenvalues are positive and some negative then the equilibrium is a saddle.

Thus, in order to understand the geometric structure of the trajectories, we need to tackle the following two fundamental questions:

1. *How many equilibria, i.e. solutions of $g = 0$, exist and where are their locations?*
2. *For each equilibrium e, what are the eigenvalues / eigenspaces of $g'(e)$?*

The main contribution of this paper is to provide "geometric" answers to these two questions for *bivariate quadratics*, which turns out to be surprisingly beautiful!

The rest of the paper is organized as follows. Section 2 lists a few genericity and simplicity assumptions for the sake of clear presentations of the main findings. Section 3 presents the main results, that is, geometric answers to the two fundamental questions identified above. Section 4 shows how to use the main results to carry out global qualitative analysis of the trajectories via combinatorial argument, without any numeric computation, from given relative positions of the solutions of f. A proof of the main result is contained in Sections 5 and 6.

2 Assumptions

Let $f = (f_1, f_2) \in \mathbb{R}[x_1, x_2]^2$ be a system of two bivariate quadratic polynomials with real coefficients. That is, there are $a_{ij}, b_{ij} \in \mathbb{R}$ such that

$$f = \begin{bmatrix} f_1 \\ f_2 \end{bmatrix} = \begin{bmatrix} \left(a_{20}x_1^2 + a_{11}x_1x_2 + a_{02}x_2^2\right) + (a_{10}x_1 + a_{01}x_2) + a_{00} \\ \left(b_{20}x_1^2 + b_{11}x_1x_2 + b_{02}x_2^2\right) + (b_{10}x_1 + b_{01}x_2) + b_{00} \end{bmatrix} \tag{3}$$

With $g = |f'|\, f'^{-1}\, f$, we will make the following assumptions.

Assumption 1 (Genericity)**.** *We assume that*

- $f = 0$ *has* 4 *simple solutions, which we will label as* r_1, r_2, r_3, r_4.
- g *is a system of two bivariate cubics with real coefficients.*
- $r_j - r_i$ *is not parallel to* $r_l - r_k$*, for any* i, j, k, l *such that* $(i, j) \neq (k, l)$.

Remark 1. *These assumptions are made to simplify the presentation of the results. Almost all bivariate quadratic systems* f *satisfy the assumptions. If the coefficients of* f *are chosen randomly, then the above assumptions hold with probability* 1.

Assumption 2 (Real solutions)**.** *We assume that all solutions of* $f = 0$ *are real such that* $r_1, r_2, r_3, r_4 \in \mathbb{R}^2$.

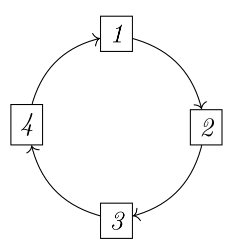
Remark 2. *Although almost all of the presented results hold for a system* f *that does not satisfy Assumption 2, we make this assumption to facilitate concrete geometrical interpretations of both the results and the corresponding proofs.*

3 Main Results

The following answers the two fundamental questions identified in the Introduction. Let f and g be defined as in Section 2 and satisfy Assumption 1 and Assumption 2. In order to state the main results compactly, we will introduce a few short-hand notations.

Notation 1.

- *Let* L_{ij} *be the line passing through* r_i *and* r_j *for* $i \neq j$*. Without loss of generality, we assume that* $i < j$ *throughout the paper.*
- $\Delta_{ij,kl} = \frac{1}{2} \begin{vmatrix} r_j - r_i & r_l - r_k \end{vmatrix}$ *is the "signed area of the triangle."*
- $\Delta_q = \Delta_{q^+q^{++},q^+q^{+++}}$ *where* $^+$ *stands for the successor on the circle in the counterclockwise direction:*



Specifically, we have

$$\Delta_1 = \Delta_{23,24}, \qquad \Delta_2 = \Delta_{34,31}, \qquad \Delta_3 = \Delta_{41,42}, \qquad \Delta_4 = \Delta_{12,13}.$$

Theorem 3 (Main Results)**.**

1. *Equilibria:*

   (a) *There are* 7 *equilibria, i.e.,* $g = 0$ *has* 7 *solutions.*

   (b) *Of these* 7 *equilibria,* 4 *are the solutions of* $f = 0$*, namely* r_1, r_2, r_3, r_4*, each called a* solution equilibrium*.*

   (c) *The other* 3 *equilibria, which we will denote as* $p_{12,34}$*,* $p_{13,24}$*, and* $p_{14,23}$*, each called a* nonsolution equilibrium*, arise as follows:* $p_{ij,kl}$ *is the intersection point of the lines* L_{ij} *and* L_{kl}*.*

2. *Eigenvalue / eigenspace at equilibrium e:*

e	*eigenvalue of* $g'(e)$	*eigenspace of* $g'(e)$
r_q	$\lambda_q = \frac{(-1)^q}{\Delta_q}\ C$	$\mathbb{R}^2$
$p_{ij,kl}$	$\lambda_{ij} = \frac{(-1)^{i+j}}{\Delta_{ij,kl}}\ C$	$\text{span}\,(r_j - r_i)$
	$\lambda_{kl} = \frac{(-1)^{k+l}}{\Delta_{kl,ij}}\ C$	$\text{span}\,(r_l - r_k)$

for some nonzero constant $C \in \mathbb{R}$*.*

The two claims of Theorem 3 are proven in Sections 5 and 6, respectively.

Remark 3. *Although an explicit expression of* C *can be written in terms of the coefficients of* f *and indexing of the roots, we will not show it here as the expression is not needed for investigating the global geometric structure of the trajectories.*

4 Geometry of the trajectories

The objectives of this section are:

- Illustrate each claim of the main results (Theorem 3) on concrete examples.
- Show that one can sketch the trajectories qualitatively via only combinatorial argument, without any numeric computation, from given relative positions of the solutions of f and the sign of the eigenvalue of g' at one solution equilibrium.

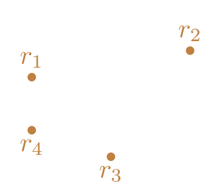
We demonstrate on two examples, which have different combinatorial structures.

Example 1. Reconsider the illustrative example from the Introduction

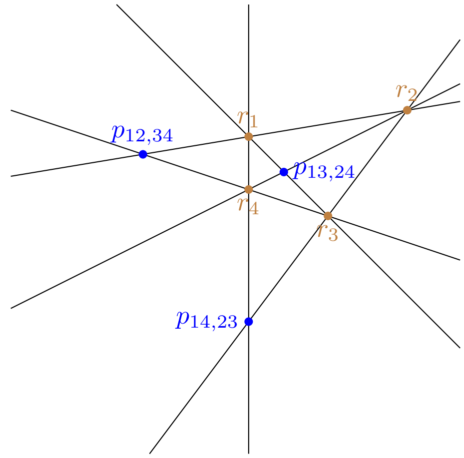
$$f = \begin{bmatrix} -11x_1^2 + 8x_1x_2 - 2x_2^2 + 63x_1 - 112 \\ -8x_1^2 + 5x_1x_2 - 8x_2^2 + 54x_1 + 54x_2 - 196 \end{bmatrix}.$$

Even though we know all the coefficients of f, we will not carry out algebraic calculations on the coefficients. Instead, we will employ geometric reasoning from the solutions of $f = 0$ to analyze (1).

1. Suppose that we are told that $f = 0$ has the following solutions.



2. By Claim 1 of Theorem 3, we determine the nonsolution equilibria geometrically by intersecting lines through pairs of solutions of $f = 0$.



3. Suppose that we are told $\lambda_1 < 0$. Using Claim 2 of Theorem 3, we can determine the signs of the eigenvalues and the direction of the corresponding eigenvectors.

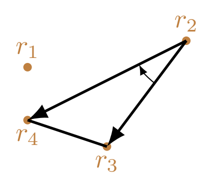
   (a) Let us first determine the sign of the constant C using $\lambda_1 < 0$. Since

$$\frac{(-1)^1}{\Delta_1} C = \lambda_1 < 0, \tag{4}$$

   we can determine the sign of C after knowing the sign of Δ_1. Recall that

$$\Delta_1 = \Delta_{23,24} = \frac{1}{2} \begin{vmatrix} r_3 - r_2 & r_4 - r_2 \end{vmatrix}$$

   which is the signed area of the following triangle:



   If the internal angle between $r_3 - r_2$ and $r_4 - r_2$ is in the counterclockwise direction, $\Delta_1 > 0$. Conversely, as in this case, if the internal angle is in the clockwise direction, $\Delta_1 < 0$. Hence, (4) yields that $C < 0$.

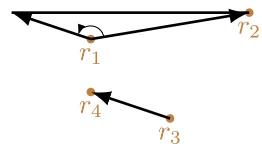
   (b) We can now determine each type of solution equilibria from the sign of Δ_q, determined similarly as above. This yields the following results:

q	$(-1)^q$	Δ_q	C	$\lambda_q = \frac{(-1)^q}{\Delta_q} C$	type
1	$-$	$-$	$-$	$-$	sink
2	$+$	$-$	$-$	$+$	source
3	$-$	$-$	$-$	$-$	sink
4	$+$	$-$	$-$	$+$	source

   (c) Next, we next analyze nonsolution equilibria. To illustrate, consider λ_{12} associated with the nonsolution equilibrium $p_{12,34}$. From Claim 2 of Theorem 3,

$$\lambda_{12} = \frac{(-1)^{1+2}}{\Delta_{12,34}} C \qquad \text{where} \qquad \Delta_{12,34} = \frac{1}{2} \begin{vmatrix} r_2 - r_1 & r_4 - r_3 \end{vmatrix}. \tag{5}$$

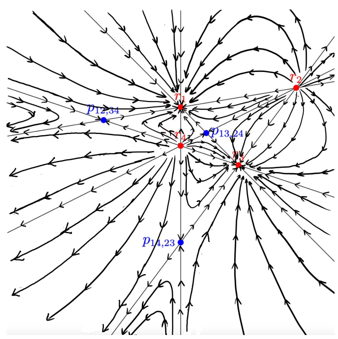
Hence, $\Delta_{12,34}$ is the signed area of the following triangle:



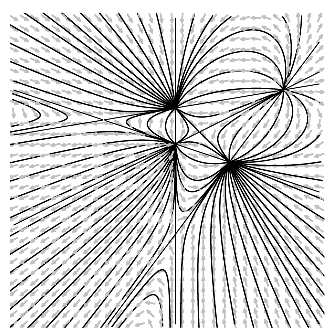
We translate the vector $r_4 - r_3$ to form the triangle. Since the direction of the interior angle is counterclockwise, $\Delta_{12,34} > 0$. Hence, (5) yields that $\lambda_{12} > 0$. Thus, $p_{12,34}$ is a source in the direction $r_2 - r_1$. Repeating this process for all of the nonsolution equilibria we obtain:

ij	kl	$(-1)^{i+j} = (-1)^{k+l}$	$\Delta_{ij,kl} = -\Delta_{kl,ij}$	C	λ_{ij}	λ_{kl}	type
12	34	$-$	$+$	$-$	$+$	$-$	saddle
13	24	$+$	$-$	$-$	$+$	$-$	saddle
14	23	$-$	$-$	$-$	$-$	$+$	saddle

4. Finally, collecting the above information on the eigenvalues/eigenspaces, we can sketch by hand the solution trajectories (left). For comparison, we also show the trajectories computed by a differential equation solver (right). Observe that the two figures match qualitatively.



Hand drawing using Theorem 3



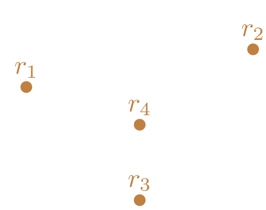
Using a differential equation solver

Example 2. Consider the bivariate quadratic system

$$f = \begin{bmatrix} x_1^2 - 17x_1x_2 + 3x_2^2 + 81x_1 + 67x_2 - 406 \\ 2x_1^2 - 23x_1x_2 + 3x_2^2 + 104x_1 + 97x_2 - 546 \end{bmatrix}.$$

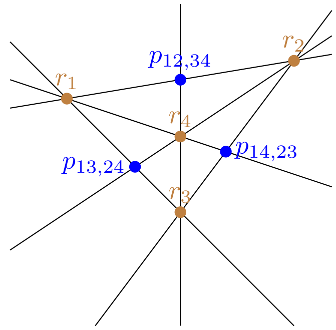
Again, let us pretend that we do not know the coefficients as we aim to analyze (1).

1. Suppose that we are told that $f = 0$ has the following solutions.



2. By Claim 1 of Theorem 3, we determine the nonsolution equilibria geometrically by intersecting lines

through pairs of solutions of $f = 0$.

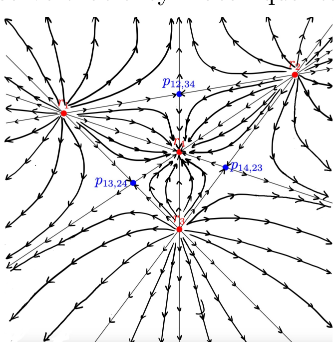


3. Suppose that $\lambda_1 > 0$. The signs of the eigenvalues and the direction of corresponding eigenvectors are computed using Claim 2 of Theorem 3.

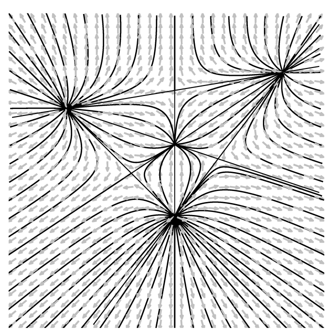
q	$(-1)^q$	Δ_q	C	$\lambda_q = \frac{(-1)^q}{\Delta_q} C$	type
1	$-$	$-$	$+$	$+$	source
2	$+$	$+$	$+$	$+$	source
3	$-$	$-$	$+$	$+$	source
4	$+$	$-$	$+$	$-$	sink

ij	kl	$(-1)^{i+j} = (-1)^{k+l}$	$\Delta_{ij,kl} = -\Delta_{kl,ij}$	C	λ_{ij}	λ_{kl}	type
12	34	$-$	$+$	$+$	$-$	$+$	saddle
13	24	$+$	$-$	$+$	$-$	$+$	saddle
14	23	$-$	$-$	$+$	$+$	$-$	saddle

4. With the above information on the eigenvalues/eigenspaces, we can sketch by hand the solution trajectories (left) and compare with trajectories computed by a differential equation solver (right). Again, observe that they match qualitatively.



Hand drawing using Theorem 3



Using a differential equation solver

5 Proof of Claim 1 of Theorem 3: Equilibria

Three subclaims, namely 1(a), 1(b), 1(c) in Theorem 3, are proven in Lemmas 4, 5, and 6, respectively. The longer proofs are divided into numbered stages for easier reading.

Lemma 4 (Claim 1(a)). *There are* 7 *solutions of* $g = 0$.

Proof.

1. Let us begin by writing g in terms of the coefficients of f. Recall that

$$g = |f'|\, f'^{-1}\, f$$

By Cramer's rule, we can simplify g as

$$g = \begin{bmatrix} \begin{vmatrix} f_1 & f_{12} \\ f_2 & f_{22} \end{vmatrix} \\ \\ \begin{vmatrix} f_{11} & f_1 \\ f_{21} & f_2 \end{vmatrix} \end{bmatrix} \in \mathbb{R}\left[x_1, x_2\right]^2$$

where f_{ij} stands for $\frac{\partial f_i}{\partial x_j}$. From (3),

$$f' = \begin{bmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{bmatrix} = \begin{bmatrix} (2a_{20}x_1 + a_{11}x_2) + (a_{10}) & (a_{11}x_1 + 2a_{02}x_2) + (a_{01}) \\ (2b_{20}x_1 + b_{11}x_2) + (b_{10}) & (b_{11}x_1 + 2b_{02}x_2) + (b_{01}) \end{bmatrix}.$$

Thus, g can be written in terms of the coefficients of f as follows.

$$g = \begin{bmatrix} \begin{vmatrix} \left(a_{20}x_1^2 + a_{11}x_1x_2 + a_{02}x_2^2\right) + (a_{10}x_1 + a_{01}x_2) + a_{00} & (a_{11}x_1 + 2a_{02}x_2) + (a_{01}) \\ \left(b_{20}x_1^2 + b_{11}x_1x_2 + b_{02}x_2^2\right) + (b_{10}x_1 + b_{01}x_2) + b_{00} & (b_{11}x_1 + 2b_{02}x_2) + (b_{01}) \end{vmatrix} \\ \\ \begin{vmatrix} (2a_{20}x_1 + a_{11}x_2) + (a_{10}) & \left(a_{20}x_1^2 + a_{11}x_1x_2 + a_{02}x_2^2\right) + (a_{10}x_1 + a_{01}x_2) + a_{00} \\ (2b_{20}x_1 + b_{11}x_2) + (b_{10}) & \left(b_{20}x_1^2 + b_{11}x_1x_2 + b_{02}x_2^2\right) + (b_{10}x_1 + b_{01}x_2) + b_{00} \end{vmatrix} \end{bmatrix}.$$

2. From the above expression, we easily infer that $\deg(g_1)$, $\deg(g_2) \leq 3$. From Assumption 1, it follows that $\deg g_1 = \deg g_2 = 3$ and that $g = 0$ has finitely many solutions. By Bezout's Theorem, e.g., see [4, Thm. 8.4.3], $g = 0$ has exactly

$$\deg(g_1) \cdot \deg(g_2) = 3 \cdot 3 = 9$$

solutions, including the solutions at infinity.

3. In order to determine the actual number of solutions of $g = 0$, it suffices to determine the number of solutions at infinity. To that end, we only need to look at the behavior of the higher degree terms, namely, the terms of degree 3. Hence, consider the "sub"-system $\hat{g}$ comprising the degree 3 terms of g, namely

$$\hat{g} = \begin{bmatrix} \begin{vmatrix} \left(a_{20}x_1^2 + a_{11}x_1x_2 + a_{02}x_2^2\right) & (a_{11}x_1 + 2a_{02}x_2) \\ \left(b_{20}x_1^2 + b_{11}x_1x_2 + b_{02}x_2^2\right) & (b_{11}x_1 + 2b_{02}x_2) \end{vmatrix} \\ \\ \begin{vmatrix} (2a_{20}x_1 + a_{11}x_2) & \left(a_{20}x_1^2 + a_{11}x_1x_2 + a_{02}x_2^2\right) \\ (2b_{20}x_1 + b_{11}x_2) & \left(b_{20}x_1^2 + b_{11}x_1x_2 + b_{02}x_2^2\right) \end{vmatrix} \end{bmatrix}.$$

After expanding and simplifying the determinant computation, we see that $\hat{g}$ has the following beautiful structure:

$$\hat{g} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \phi$$

where

$$\phi = (a_{20}b_{11} - a_{11}b_{20})x_1^2 + 2(a_{20}b_{02} - a_{02}b_{20})x_1x_2 + (a_{11}b_{02} - a_{02}b_{11})x_2^2.$$

4. Since solutions of $g = 0$ at infinity correspond with solutions of $\hat{g} = 0$ on $\mathbb{P}^1$, we only need to consider $\phi = 0$ in $\mathbb{P}^1$. From Assumption 1, we have that $\deg \phi = 2$. Thus, it follows from the Fundamental Theorem of Algebra, e.g., see [4, Thm. 5.1.1], that $g = 0$ has 2 solutions at infinity. Thus, the actual number of solutions to $g = 0$ is $9 - 2 = 7$.

□

Lemma 5 (Claim 1(b))**.** *If e is such that $f(e) = 0$, then $g(e) = 0$.*

Proof. This follows immediately from

$$g(e) = \begin{bmatrix} \begin{vmatrix} f_1(e) & f_{12}(e) \\ f_2(e) & f_{22}(e) \end{vmatrix} \\ \\ \begin{vmatrix} f_{11}(e) & f_1(e) \\ f_{21}(e) & f_2(e) \end{vmatrix} \end{bmatrix} = \begin{bmatrix} \begin{vmatrix} 0 & f_{12}(e) \\ 0 & f_{22}(e) \end{vmatrix} \\ \\ \begin{vmatrix} f_{11}(e) & 0 \\ f_{21}(e) & 0 \end{vmatrix} \end{bmatrix} = 0.$$

□

Lemma 6 (Claim 1(c))**.** *If $p = p_{ij,kl}$ is the intersection of two lines L_{ij} and L_{kl} passing through pairs of solutions of $f = 0$, then $g(p) = 0$.*

Proof.

1. Let us begin by studying the relation between f and f' along the line L_{ij}. For this, consider the parametrization of L_{ij} given by
$$h_{ij}(s) = r_i + s(r_j - r_i),$$
and define
$$\bar{f} = f(h_{ij}(s)) \in \mathbb{R}[s]^2$$
where s is an indeterminate. Note $\bar{f}$ is a system of two univariate polynomials with $\deg_s \bar{f} \leq 2$ such that
$$\bar{f}(0) = f(r_i) = 0 \quad \text{and} \quad \bar{f}(1) = f(r_j) = 0.$$
Hence, we know that $\bar{f}$ has the form
$$\bar{f} = s(s-1)c_{ij} = f(h_{ij}(s)) \tag{6}$$
for some vector $c_{ij} \in \mathbb{R}^2$. By differentiating $\bar{f}$ with respect to s, we have
$$\frac{d\bar{f}}{ds} = (2s-1)c_{ij} = f'(h_{ij}(s))\, h'_{ij}(s) = f'(h_{ij}(s))\,(r_j - r_i). \tag{7}$$
By eliminating c_{ij} from (7) using (6), we have
$$f'(h_{ij}(s))\,(r_j - r_i) = \frac{d\bar{f}}{ds} = \frac{2s-1}{s(s-1)}\bar{f} = \frac{2s-1}{s(s-1)} f(h_{ij}(s)). \tag{8}$$
By repeating the same process over L_{kl}, we obtain
$$f'(h_{kl}(s))\,(r_l - r_k) = \frac{2s-1}{s(s-1)} f(h_{kl}(s)) \tag{9}$$

2. Consider (8) and (9) at the intersection point p in terms of s_{ij} and s_{kl}

$$h_{ij}(s_{ij}) = r_i + s_{ij}(r_j - r_i) = p = r_k + s_{kl}(r_l - r_k) = h_{kl}(s_{kl}). \tag{10}$$

By instantiating (8) at $s = s_{ij}$ and (9) at $s = s_{kl}$, we have

$$f'(p)(r_j - r_i) = \frac{2s_{ij} - 1}{s_{ij}(s_{ij} - 1)} f(p)$$
$$f'(p)(r_l - r_k) = \frac{2s_{kl} - 1}{s_{kl}(s_{kl} - 1)} f(p).$$

Rewriting in matrix form, this yields

$$f'(p) \begin{bmatrix} r_j - r_i & r_l - r_k \end{bmatrix} = f(p) \begin{bmatrix} \frac{2s_{ij}-1}{s_{ij}(s_{ij}-1)} & \frac{2s_{kl}-1}{s_{kl}(s_{kl}-1)} \end{bmatrix}$$

By solving for $f'(p)$, we obtain

$$f'(p) = f(p)\, u \tag{11}$$

where

$$u = \begin{bmatrix} u_1 & u_2 \end{bmatrix} = \begin{bmatrix} \frac{2s_{ij}-1}{s_{ij}(s_{ij}-1)} & \frac{2s_{kl}-1}{s_{kl}(s_{kl}-1)} \end{bmatrix} \begin{bmatrix} r_j - r_i & r_l - r_k \end{bmatrix}^{-1} \in \mathbb{R}^{1\times 2}$$

Note that u is well-defined due to Assumption 1. In particular, writing (11) explicitly yields

$$\begin{bmatrix} f_{11}(p) & f_{12}(p) \\ f_{21}(p) & f_{22}(p) \end{bmatrix} = \begin{bmatrix} f_1(p) \\ f_2(p) \end{bmatrix} \begin{bmatrix} u_1 & u_2 \end{bmatrix} = \begin{bmatrix} f_1(p)\, u_1 & f_1(p)\, u_2 \\ f_2(p)\, u_1 & f_2(p)\, u_2 \end{bmatrix}.$$

3. The final observation is that

$$g(p) = \begin{bmatrix} \begin{vmatrix} f_1(p) & f_{12}(p) \\ f_2(p) & f_{22}(p) \end{vmatrix} \\ \\ \begin{vmatrix} f_{11}(p) & f_1(p) \\ f_{21}(p) & f_2(p) \end{vmatrix} \end{bmatrix} = \begin{bmatrix} \begin{vmatrix} f_1(p) & f_1(p)\, u_2 \\ f_2(p) & f_2(p)\, u_2 \end{vmatrix} \\ \\ \begin{vmatrix} f_1(p)\, u_1 & f_1(p) \\ f_2(p)\, u_1 & f_2(p) \end{vmatrix} \end{bmatrix} = 0.$$

□

6 Proof of Claim 2 of Theorem 3: Eigenvalues / eigenspaces

We start with two lemmas before proceeding to prove Claim 2 of Theorem 3.

Lemma 7. *Let e be a solution equilibrium, that is, $g(e) = f(e) = 0$. The matrix $g'(e)$ has one and only one eigenvalue with the corresponding eigenspace being $\mathbb{R}^2$.*

Proof. For $f_{ij} = \frac{\partial f_i}{\partial x_j}$ and $f_{ijk} = \frac{\partial^2 f_i}{\partial x_j \partial x_k}$, we have

$$g' = \begin{bmatrix} \begin{vmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{vmatrix} + \begin{vmatrix} f_1 & f_{121} \\ f_2 & f_{221} \end{vmatrix} & \begin{vmatrix} f_{12} & f_{12} \\ f_{22} & f_{22} \end{vmatrix} + \begin{vmatrix} f_1 & f_{122} \\ f_2 & f_{222} \end{vmatrix} \\ \begin{vmatrix} f_{111} & f_1 \\ f_{211} & f_2 \end{vmatrix} + \begin{vmatrix} f_{11} & f_{11} \\ f_{21} & f_{21} \end{vmatrix} & \begin{vmatrix} f_{112} & f_1 \\ f_{212} & f_2 \end{vmatrix} + \begin{vmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{vmatrix} \end{bmatrix}$$

$$= \begin{bmatrix} |f'| + \begin{vmatrix} f_1 & f_{121} \\ f_2 & f_{221} \end{vmatrix} & \begin{vmatrix} f_1 & f_{122} \\ f_2 & f_{222} \end{vmatrix} \\ \begin{vmatrix} f_{111} & f_1 \\ f_{211} & f_2 \end{vmatrix} & \begin{vmatrix} f_{112} & f_1 \\ f_{212} & f_2 \end{vmatrix} + |f'| \end{bmatrix}.$$

Hence,

$$g'(e) = \begin{bmatrix} |f'(e)| & 0 \\ 0 & |f'(e)| \end{bmatrix} = |f'(e)|\, I$$

where I is the 2×2 identity matrix. Thus, $g'(e)$ has only eigenvalue, namely $|f'(e)|$, and the corresponding eigenspace is $\mathbb{R}^2$. □

Lemma 8. *Following the notation from Lemma 6, if* $p_{ij,kl} = h_{ij}(s_{ij}) = h_{kl}(s_{kl})$*, then*

$$s_{ij} = \frac{\Delta_{kl,ki}}{\Delta_{ij,kl}}$$

Proof. Rearranging (10) yields

$$(r_j - r_i)\, s_{ij} + (r_k - r_l)\, s_{kl} = r_k - r_i.$$

Rewriting it as a system of equations, we have

$$\begin{bmatrix} r_j - r_i & r_k - r_l \end{bmatrix} \begin{bmatrix} s_{ij} \\ s_{kl} \end{bmatrix} = r_k - r_i$$

Using Cramer's rule and rewriting using properties of determinants, we obtain

$$s_{ij} = \frac{\begin{vmatrix} r_k - r_i & r_k - r_l \end{vmatrix}}{\begin{vmatrix} r_j - r_i & r_k - r_l \end{vmatrix}} = -\frac{\begin{vmatrix} r_i - r_k & r_l - r_k \end{vmatrix}}{\begin{vmatrix} r_j - r_i & r_l - r_k \end{vmatrix}} = \frac{\begin{vmatrix} r_l - r_k & r_i - r_k \end{vmatrix}}{\begin{vmatrix} r_j - r_i & r_l - r_k \end{vmatrix}} = \frac{\Delta_{kl,ki}}{\Delta_{ij,kl}}.$$

□

With Lemma 7 and Lemma 8, we now complete the proof of the remaining parts.

Proof of Claim 2 of Theorem 3.

1. We will first derive relations among λ_{ij}, λ_i, and λ_j. For this, observe that $g(h_{ij}(s))$ is a system of two univariate polynomials with degree at most 3 with

$$\begin{aligned} g(h_{ij}(0)) &= g(r_i) = 0, \\ g(h_{ij}(1)) &= g(r_j) = 0, \\ g(h_{ij}(s_{ij})) &= g(p_{ij,kl}) = 0. \end{aligned}$$

From Assumption 1, we have that $p_{ij,kl}$ is neither r_i nor r_j, that is, $s_{ij} \neq 0$ and $s_{ij} \neq 1$. Hence, there is a vector $v_{ij} \in \mathbb{R}^2$ such that

$$g(h_{ij}(s)) = (s-0)(s-1)(s-s_{ij})\, v_{ij}. \tag{12}$$

Differentiating both sides with respect to s yields

$$g'(h_{ij}(s))(r_j - r_i) = ((s-1)(s-s_{ij}) + (s-0)(s-s_{ij}) + (s-0)(s-1))\, v_{ij}.$$

Evaluating at $s = 0$, $s = 1$ and $s = s_{ij}$ provides

$$g'(r_i)(r_j - r_i) = (0-1)(0-s_{ij})v_{ij} = s_{ij}v_{ij}, \tag{13}$$

$$g'(r_j)(r_j - r_i) = (1-0)(1-s_{ij})v_{ij} = (1-s_{ij})v_{ij}, \tag{14}$$
$$g'(p_{ij,kl})(r_j - r_i) = (s_{ij}-0)(s_{ij}-1)v_{ij} = s_{ij}(s_{ij}-1)v_{ij}. \tag{15}$$

From Lemma 7, we know that $g'(r_i)$ and $g'(r_j)$ each have exactly one eigenvalue, λ_i and λ_j, respectively, with eigenspace $\mathbb{R}^2$. Thus, we have

$$g'(r_i)(r_j - r_i) = \lambda_i (r_j - r_i), \tag{16}$$
$$g'(r_j)(r_j - r_i) = \lambda_j (r_j - r_i). \tag{17}$$

Combining (13), (14), (16), and (17), we have

$$\lambda_i (r_j - r_i) = s_{ij} v_{ij},$$
$$\lambda_j (r_j - r_i) = (1-s_{ij}) v_{ij},$$

which, combined with (15), yields

$$g'(p_{ij,kl})(r_j - r_i) = (s_{ij}-1)\lambda_i (r_j - r_i)$$
$$g'(p_{ij,kl})(r_j - r_i) = -s_{ij}\lambda_j (r_j - r_i)$$

This shows that $r_j - r_i$ is an eigenvector of $g'(p_{ij,kl})$ with eigenvalue

$$\lambda_{ij} = (s_{ij}-1)\lambda_i = -s_{ij}\lambda_j. \tag{18}$$

From Lemma 8, we have

$$\lambda_{ij} = \left(\frac{\Delta_{kl,ki}}{\Delta_{ij,kl}} - 1\right)\lambda_i = -\frac{\Delta_{kl,ki}}{\Delta_{ij,kl}}\lambda_j.$$

Clearing the denominator, which is nonzero by Assumption 1, and introducing a new variable $E_{ij,kl}$ yields

$$\Delta_{ij,kl}\ \lambda_{ij} = (\Delta_{kl,ki} - \Delta_{ij,kl})\ \lambda_i = -\Delta_{kl,ki}\ \lambda_j = E_{ij,kl}.$$

Note that

$$\Delta_{kl,ki} - \Delta_{ij,kl} = -\Delta_{ki,kl} - \Delta_{ij,kl} = -\Delta_{kj,kl} = \Delta_{jk,kl} = \Delta_{jk,jl}$$

yields the following relations among λ_{ij}, λ_i, and λ_j:

$$\Delta_{ij,kl}\ \lambda_{ij} = \Delta_{jk,jl}\ \lambda_i = -\Delta_{kl,ki}\ \lambda_j = E_{ij,kl}. \tag{19}$$

Looping over the choices of i, j, k, and l yields a linear system which could be solved mechanically and directly. Rather than take such an approach, we dig a little deeper into the relations to provide a systematic analysis of these equations.

2. By instantiating the relation (19), we have

$$\begin{array}{ccccccc}
\Delta_{ij,kl}\ \lambda_{ij} & = & \Delta_{jk,jl}\ \lambda_i & = & -\Delta_{kl,ki}\ \lambda_j & = & E_{ij,kl} \\
\hline
\Delta_{12,34}\ \lambda_{12} & = & \Delta_{23,24}\ \lambda_1 & = & -\Delta_{34,31}\ \lambda_2 & = & E_{12,34} \\
\Delta_{34,12}\ \lambda_{34} & = & \Delta_{41,42}\ \lambda_3 & = & -\Delta_{12,13}\ \lambda_4 & = & E_{34,12} \\
\Delta_{13,24}\ \lambda_{13} & = & \Delta_{32,34}\ \lambda_1 & = & -\Delta_{24,21}\ \lambda_3 & = & E_{13,24} \\
\Delta_{24,13}\ \lambda_{24} & = & \Delta_{41,43}\ \lambda_2 & = & -\Delta_{13,12}\ \lambda_4 & = & E_{24,13} \\
\Delta_{14,23}\ \lambda_{14} & = & \Delta_{42,43}\ \lambda_1 & = & -\Delta_{23,21}\ \lambda_4 & = & E_{14,23} \\
\Delta_{23,14}\ \lambda_{23} & = & \Delta_{31,34}\ \lambda_2 & = & -\Delta_{14,12}\ \lambda_3 & = & E_{23,14}.
\end{array} \tag{20}$$

Let us rewrite in terms of short-hand notation introduced in Section 3. For instance, $\Delta_{23,24}$ and $\Delta_{32,34}$ can both be written in terms of Δ_1, namely

$$\Delta_{23,24} = \Delta_1 \qquad \text{and} \qquad \Delta_{32,34} = \Delta_{32,24} = -\Delta_{23,24} = -\Delta_1.$$

By carrying out such rewriting for the rest, (20) becomes

$$
\begin{array}{rcrcrcl}
\Delta_{12,34}\ \lambda_{12} & = & \Delta_1\lambda_1 & = & -\Delta_2\ \lambda_2 & = & E_{12,34} \\
\Delta_{34,12}\ \lambda_{34} & = & \Delta_3\ \lambda_3 & = & -\Delta_4\ \lambda_4 & = & E_{34,12} \\
\Delta_{13,24}\ \lambda_{13} & = & -\Delta_1\ \lambda_1 & = & -\Delta_3\ \lambda_3 & = & E_{13,24} \\
\Delta_{24,13}\ \lambda_{24} & = & \Delta_2\ \lambda_2 & = & \Delta_4\ \lambda_4 & = & E_{24,13} \\
\Delta_{14,23}\ \lambda_{14} & = & \Delta_1\ \lambda_1 & = & -\Delta_4\ \lambda_4 & = & E_{14,23} \\
\Delta_{23,14}\ \lambda_{23} & = & -\Delta_2\ \lambda_2 & = & \Delta_4\ \lambda_3 & = & E_{23,14}
\end{array}
\tag{21}
$$

3. Let us investigate relations among the E's. From (21), we obtain

$$
\begin{array}{rcrcrcr}
\Delta_1\lambda_1 & = & E_{12,34} & = & -E_{13,24} & = & E_{14,23} \\
\Delta_2\lambda_2 & = & E_{24,13} & = & -E_{23,14} & = & -E_{12,34} \\
\Delta_3\lambda_3 & = & E_{34,12} & = & -E_{13,24} & = & E_{23,14} \\
\Delta_4\lambda_4 & = & -E_{34,12} & = & E_{24,13} & = & -E_{14,23}
\end{array}
$$

For just the E's, this provides 8 homogeneous equations in the 6 unknown E's. While this may appear to be overdetermined, there are only 5 linearly independent conditions resulting in a one-dimensional solution set which can be parameterized by a parameter C as follows:

$$
\begin{array}{rcr}
E_{12,34} & = & -C \\
E_{34,12} & = & -C \\
E_{13,24} & = & C \\
E_{24,13} & = & C \\
E_{14,23} & = & -C \\
E_{23,14} & = & -C
\end{array}
$$

Using equations involving λ_q, the value of C can be determined, e.g., $\Delta_1\lambda_1 = E_{12,34} = -C$. Thus, (21) becomes

$$
\begin{array}{rcrcrcr}
\Delta_{12,34}\ \lambda_{12} & = & \Delta_1\lambda_1 & = & -\Delta_2\ \lambda_2 & = & -C \\
\Delta_{34,12}\ \lambda_{34} & = & \Delta_3\ \lambda_3 & = & -\Delta_4\ \lambda_4 & = & -C \\
\Delta_{13,24}\ \lambda_{13} & = & -\Delta_1\ \lambda_1 & = & -\Delta_3\ \lambda_3 & = & C \\
\Delta_{24,13}\ \lambda_{24} & = & \Delta_2\ \lambda_2 & = & \Delta_4\ \lambda_4 & = & C \\
\Delta_{14,23}\ \lambda_{14} & = & \Delta_1\ \lambda_1 & = & -\Delta_4\ \lambda_4 & = & -C \\
\Delta_{23,14}\ \lambda_{23} & = & -\Delta_2\ \lambda_2 & = & \Delta_4\ \lambda_3 & = & -C
\end{array}
$$

so that

$$
\begin{array}{llll}
\lambda_1 = \frac{-1}{\Delta_1}C, & \lambda_2 = \frac{1}{\Delta_2}C, & \lambda_3 = \frac{-1}{\Delta_3}C, & \lambda_4 = \frac{1}{\Delta_4}C, \\
\lambda_{12} = \frac{-1}{\Delta_{12,34}}C, & \lambda_{13} = \frac{1}{\Delta_{13,24}}C, & \lambda_{14} = \frac{-1}{\Delta_{14,23}}C, & \\
\lambda_{34} = \frac{-1}{\Delta_{34,12}}C, & \lambda_{24} = \frac{1}{\Delta_{24,13}}C, & \lambda_{23} = \frac{-1}{\Delta_{23,14}}C. &
\end{array}
$$

4. We can rewrite the above solution more compactly as follows. For $1 \le q \le 4$ and $1 \le i < j \le 4$, we have

$$
\lambda_q = \frac{(-1)^q}{\Delta_q}C \qquad \text{and} \qquad \lambda_{ij} = \frac{(-1)^{i+j}}{\Delta_{ij,kl}}C.
$$

□

Acknowledgments

JDH and CH were supported in part by National Science Foundation grant CCF 2331400. HH and FPC were supported in part by National Science Foundation grants CCF 2331401. JDH also acknowledges support from the Simons Foundation SFM-00005696 and the Robert and Sara Lumpkins Collegiate Professorship. HH also acknowledges support from NSF CCF 2212461.

References

[1] Diener, I, On the global convergence of path-following methods to determine all solutions to a system of nonlinear equations, *Mathematical Programming*, 39, 181–188, 1987.

[2] Hartman, P., *Ordinary Differential Equations*, Classics in Applied Mathematics, 38, SIAM, Philadelphia, PA, 2002.

[3] Motyka, M.A. and Reiter, C.A., Chaos and Newton's method on systems, *Computers & Graphics*, 14(1), 131–134, 1990.

[4] Sommese, A.J. and Wampler, C.W., *The Numerical Solution of Systems of Polynomials Arising in Engineering and Science*, World Sci. Publ., Hackensack, NJ, 2005.

[5] Ypma, T.J., Historical development of the Newton-Raphson method, *SIAM Review*, 37(4), 530–551, 1995.