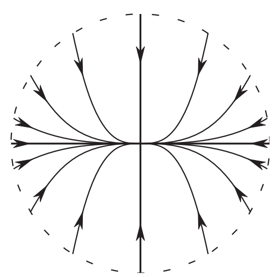
Conjugate Phase Portraits of Linear Systems

Patrick D. McSwiggen and Kenneth R. Meyer

1. INTRODUCTION. A standard topic in a first course in the geometric theory of differential equations is the classification of phase portraits of linear systems. One considers equations of the form

$$\dot{x} = Ax \tag{1}$$

where $x \in \mathbb{R}^2$, A is a 2×2 real constant matrix, and the dot denotes the derivative with respect to $t \in \mathbb{R}$, $\dot{} = d/dt$. The solution of (1) such that $x = x_0$ when $t = 0$ is $x(t, x_0) = e^{At}x_0$. A *trajectory* of a system $\dot{x} = Ax$ is a solution of the system, i.e., a map $\mathbb{R} \to \mathbb{R}^2 : t \mapsto e^{At}x_0$, whereas an *orbit* is the oriented geometric curve $\{e^{At}x_0 \in \mathbb{R}^2 : t \in \mathbb{R}\}$, where the orientation is given by increasing t. The *phase portrait* of a system (1) is the totality of all its orbits in $\mathbb{R}^2$, and is illustrated graphically by drawing a few judicially chosen orbits. For example we draw the picture



to illustrate the phase portrait of the system

$$\dot{x}_1 = -2x_1$$
$$\dot{x}_2 = -3x_2.$$

The solution through (ξ_1, ξ_2) at $t = 0$ is $x_1(t) = e^{-2t}\xi_1$, $x_2(t) = e^{-3t}\xi_2$, so all solutions are decreasing as $t \to \infty$, the coordinate axes are orbits, and the other orbits are in the family of cusps $x_1^3 = \gamma x_2^2$, $\gamma = \xi_1^3/\xi_2^2$. The orientation or sense of the trajectory is shown by an arrow and indicates increasing t in phase portraits.

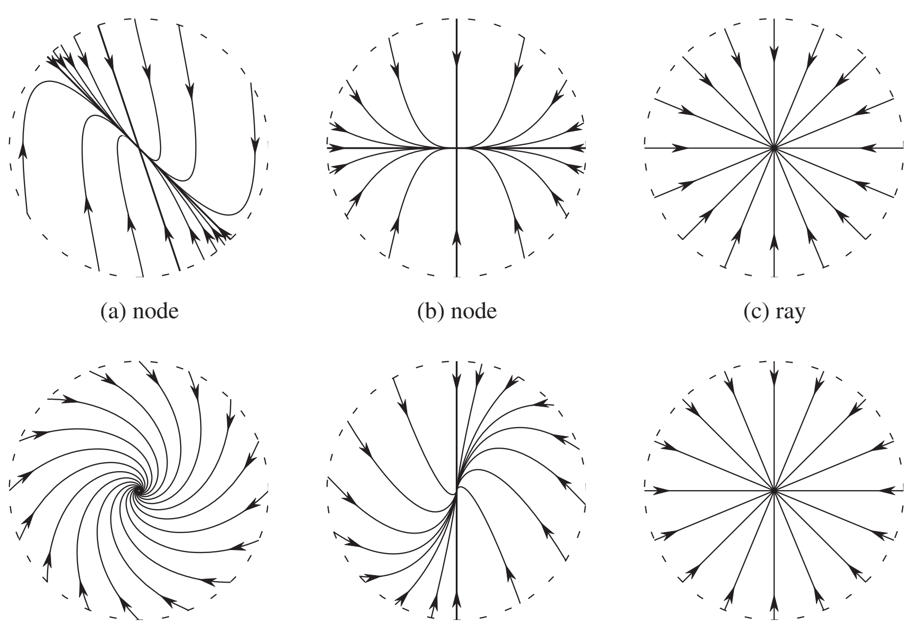
Phase portraits are discussed in elementary texts like Blanchard, Devaney, and Hall [**1**] and Boyce and DiPrima [**2**] and in advanced texts like Hubbard and West [**4**] and Robinson [**6**]. In these texts, it is emphasized that the geometry of these portraits depends heavily on the eigenvalues and eigenvectors of the coefficient matrix. The portraits depend on the signs and the magnitudes of the real parts of the eigenvalues, the existence or nonexistence of imaginary parts of the eigenvalues, and the dimensions of the eigenspaces.

Klein's Erlangen project defines a geometry by a group of transformations. For us the geometry of the phase portraits is defined by conjugacies. A *conjugacy* of two systems of the form (1) is a homeomorphism of $\mathbb{R}^2$ which takes the trajectories of one system onto the trajectories of the other while preserving the time parameter, t. If such a conjugacy exists, then we say the two systems are *conjugate*. Here we restrict the group of conjugacies to different smoothness classes (topological, Hölder, Lipschitz, linear) and investigate what these classes tell us about the geometry of the phase portraits and about the coefficient matrices. This hierarchy of smoothness reveals subtle differences in the properties of linear systems as a whole.

Looked at from the opposite point of view, this treatise is a case study on the invariants of homeomorphisms of different levels of smoothness. In general topology one studies the invariants of continuous homeomorphisms and in differential topology one studies the invariants of smooth diffeomorphism. But, what about Hölder or Lipschitz homeomorphisms? As one goes up the hierarchy of smoothness one discerns finer and finer geometric structure.

We completely characterize the conjugacy of n-dimensional hyperbolic systems, i.e., systems for which all the eigenvalues of the coefficient matrix have nonzero real part. But first, in the next section six representative two-dimensional examples are given to illustrate the various possible planar phase portraits, and the question of the level of smoothness of the conjugacies between them is completely answered. It is left to the final section to completely resolve the general n-dimensional hyperbolic case. It turns out that with one exception the general case is the same as the planar examples. The exceptional case contains a $2k$-dimensional Jordan block (see Theorem 3.1) with complex eigenvalues, where $k \geq 2$.

2. PLANAR EXAMPLES. Figure 1 contains six examples of the phase portraits of linear differential equations in the plane, i.e., equations of the form (1) where $x \in \mathbb{R}^2$ and A is a 2×2 real constant matrix. Since the systems are two-dimensional, the phase portraits are planar. The n-dimensional case will be discussed in the following section.



(a) node (b) node (c) ray

(d) focus (e) one-tangent node (f) ray

Figure 1. Phase portraits of selected linear systems.

The solutions of equations (1), and hence the phase portraits in Figure 1, depend on the eigenstructure (i.e., eigenvalues and eigenvectors) of the matrix A. The six systems of equations which give rise to the phase portraits in Figure 1 are

$$
\begin{array}{llllll}
\text{(a)} & \begin{aligned}\dot u_1 &= v_1\\ \dot v_1 &= -6u_1 - 5v_1\end{aligned} & \text{(b)} & \begin{aligned}\dot u_2 &= -2u_2\\ \dot v_2 &= -3v_2\end{aligned} & \text{(c)} & \begin{aligned}\dot u_3 &= -2u_3\\ \dot v_3 &= -2v_3\end{aligned}\\
\text{(d)} & \begin{aligned}\dot u_4 &= -2u_4 + v_4\\ \dot v_4 &= -u_4 - 2v_4\end{aligned} & \text{(e)} & \begin{aligned}\dot u_5 &= -2u_5\\ \dot v_5 &= u_5 - 2v_5\end{aligned} & \text{(f)} & \begin{aligned}\dot u_6 &= -u_6\\ \dot v_6 &= -v_6\end{aligned}
\end{array}
\tag{2}
$$

For each example the unknown is $x_k = (u_k, v_k)^T$, the corresponding coefficient matrices are

$$
\begin{array}{lll}
\text{(a) } A_1 = \begin{bmatrix} 0 & 1 \\ -6 & -5 \end{bmatrix} & \text{(b) } A_2 = \begin{bmatrix} -2 & 0 \\ 0 & -3 \end{bmatrix} & \text{(c) } A_3 = \begin{bmatrix} -2 & 0 \\ 0 & -2 \end{bmatrix}\\
\text{(d) } A_4 = \begin{bmatrix} -2 & 1 \\ -1 & -2 \end{bmatrix} & \text{(e) } A_5 = \begin{bmatrix} -2 & 0 \\ 1 & -2 \end{bmatrix} & \text{(f) } A_6 = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}
\end{array}
\tag{3}
$$

and their eigenvalues are

$$
\begin{array}{lll}
\text{(a) } -2, -3 & \text{(b) } -2, -3 & \text{(c) } -2, -2\\
\text{(d) } -2 \pm i & \text{(e) } -2, -2 & \text{(f) } -1, -1
\end{array}
\tag{4}
$$

The matrices in (3a) and (3b) have the same distinct eigenvalues and so are similar. The matrix in (3c) is diagonal, so the eigenspace associated to -2 is all of $\mathbb{R}^2$, whereas the matrix in (3e) has an off-diagonal term, so its eigenspace is one dimensional, specifically span$\{(0, 1)^T\}$. The matrices in (3c) and (3f) are both diagonal with repeated eigenvalues -2 and -1, respectively. The matrix in (3d) has a single conjugate pair of complex eigenvalues.

In all of these examples the real parts of the eigenvalues are negative, so the solutions are exponentially decreasing, as indicated by the arrows in the figures. One says that they are all *sinks*. One obtains a *source*, where all the real parts of the eigenvalues are positive, by the time reversal $t \to -t$ and reversing the arrows in the phase portraits. The real part of the eigenvalue determines the rate at which solutions approach the origin—the *exponential rate of decay*. All solutions of (2c) approach the origin at the same rate and faster than the solutions of (2f), as the placement of arrows in Figures 1c and 1f is meant to imply, whereas the solutions of (2b) approach the origin more rapidly along the vertical axis than along the horizontal axis. The rate at which solutions rotate around the origin is determined by the imaginary parts of the eigenvalues.

To the eye these figures have both qualitative differences and similarities, and it is these features we wish to discuss. We will present several different definitions of conjugate phase portraits for n-dimensional systems. The different definitions distinguish different geometric features of the portraits and the properties of the eigenvalues and eigenvectors of the matrix A.

Consider two 2-dimensional systems $\dot x = Ax$ and $\dot y = By$ with their corresponding phase portraits.

- The two systems are *topologically conjugate* if there is a homeomorphism $\phi : \mathbb{R}^2 \to \mathbb{R}^2$ that carries the trajectories of the first system onto the trajectories of the second and preserves the time parameter, t. That is, $\phi(e^{At}x_0) = e^{Bt}\phi(x_0)$, for all $x_0 \in \mathbb{R}^2$, $t \in \mathbb{R}$.

- If ϕ and ϕ^{-1} are Hölder continuous with Hölder exponent α $(0 < \alpha < 1)$, then the two systems are *α-Hölder conjugate*. Here we say that ϕ is Hölder continuous with exponent α if for every compact set $C \subset \mathbb{R}^2$ there is a constant K such that $\|\phi(u) - \phi(v)\| \le K\|u - v\|^\alpha$ for all $u, v \in C$.
- If ϕ and ϕ^{-1} are Hölder continuous for all exponents α satisfying $0 < \alpha < 1$ then we will refer to the two systems as simply *Hölder conjugate*.
- If ϕ and ϕ^{-1} are Lipschitz continuous then the two systems are *Lipschitz conjugate*. Here we say that ϕ is Lipschitz continuous if for every compact set $C \subset \mathbb{R}^2$ there is a constant K such that $\|\phi(u) - \phi(v)\| \le K\|u - v\|$ for all $u, v \in C$.
- If ϕ and ϕ^{-1} are C^1 then the two systems are C^1 *conjugate*.
- If ϕ and ϕ^{-1} are linear then the two systems are *linearly conjugate*.
- In all these cases we say ϕ is a *conjugacy*.

Lemma 2.1. *The two systems are linearly conjugate if and only if A and B are similar, i.e., there exists a nonsingular matrix P such that $B = PAP^{-1}$. If the two systems are C^1 conjugate then they are linearly conjugate.*

Proof. Let the two systems be linearly conjugate, and let $\phi(x) = Px$ be the conjugacy. Differentiate $\phi(e^{At}x_0) = Pe^{At}x_0 = e^{Bt}\phi(x_0) = e^{Bt}Px_0$ with respect to t and then set $t = 0$ to get $PAx_0 = BPx_0$ for all x_0. Hence $PA = BP$ and A and B are similar. Conversely, if $PA = BP$ then $Pe^{At} = e^{Bt}P$, so $\phi(x) = Px$ is a conjugacy.

Let the two systems be C^1 conjugate. Differentiating $\phi(e^{At}x_0) = e^{Bt}\phi(x_0)$ with respect to x_0 and setting $x_0 = 0$ produces $Pe^{At} = e^{Bt}P$, where P is the Jacobian of ϕ at the origin. Thus, A and B are similar. ■

The order of the definitions above is from the weaker to the stronger, i.e., $C^1 \Longleftrightarrow$ linear $\Rightarrow$ Lipschitz $\Rightarrow$ Hölder $\Rightarrow$ topological conjugacy. These differences in smoothness detect the differences in the eigenstructure as illustrated by Proposition 2.1 below.

There is also a weaker notion of equivalence between two systems. The two systems are *topologically equivalent* if there is a homeomorphism $\phi : \mathbb{R}^2 \to \mathbb{R}^2$ which carries the orbits of the first system onto the orbits of the second and preserves the sense of the orbits, and in this case ϕ is called an *equivalence*. The equivalence is Hölder, Lipschitz, or linear as ϕ and ϕ^{-1} are Hölder, Lipschitz, or linear, respectively. A conjugacy preserves trajectories, whereas an equivalence preserves orbits.

If ϕ is a conjugacy then it is automatically an equivalence, but the converse is not necessarily true. The identity map takes the orbits of system (2c) to the orbits of (2f) and so these systems are linearly equivalent, but as we shall see they are not even Hölder conjugate. We will primarily restrict our attention here to conjugacies. However, since equivalences are more general, when we wish to prove the nonexistence of a certain conjugacy, it will suffice to prove the absence of a corresponding equivalence.

Proposition 2.1. *All systems in* (2) *with phase portraits in Figure 1 are topologically conjugate. More specifically,*

- *systems* (2a) *and* (2b)*, with phase portraits in Figures 1a and 1b, are linearly conjugate, i.e., the matrices A_1 and A_2 are similar,*
- *systems* (2c) *and* (2d)*, with phase portraits in Figures 1c and 1d, where the systems differ only in their rate of rotation, are Lipschitz conjugate,*
- *systems* (2c) *and* (2e)*, with phase portraits in Figures 1c and 1e, where the systems have the same eigenvalues, but eigenspaces of different dimension, are Hölder conjugate,*

- *systems* (2c) *and* (2f), *with phase portraits in Figures 1c and 1f, where the exponential rates of decay differ, are* $\frac{1}{2}$*-Hölder conjugate. (Likewise, systems* (2c) *and* (2b), *with phase portraits in Figures 1c and 1b, are* $\frac{2}{3}$*-Hölder conjugate.)*

Moreover, in each case these are the best possible for the conjugacies under consideration.

One of the key features of the above proposition, and the point of the chosen examples, is the presence of a hierarchy within the building blocks of the coefficient matrix. We think of a matrix as being constructed from its eigenvalues (real and imaginary parts treated separately), the Jordan block structure for each eigenvalue (see Theorem 3.1), and the physical orientations of the eigenvectors and generalized eigenspaces in space. Clearly, the specific eigenvectors and placement of generalized eigenspaces can be modified using a linear change of variable—the strongest conjugacy. This sort of change of variables is so automatic that the actual eigenvectors seem more an artifact of how we chose our axes than an intrinsic part of the system. A linear change of variables does not affect the other properties of the system, making them seem perhaps more inherent or at least more tightly bound to what is the essence of the system. However, the above examples suggest that there is a further gradation within these other properties—some properties more tenuous and easily disrupted than others.

The next weakest conjugacy is Lipschitz conjugacy which ignores the rotational term and transforms a focus to a ray as long as they have the same exponential rate of decay. Thus a Lipschitz conjugacy does not see the rotation. However, a Lipschitz conjugacy cannot modify the off-diagonal terms in a Jordan block, whereas a Hölder conjugacy can. A Hölder conjugacy ignores the off-diagonal terms in the matrix, which are what give rise to the tangencies in Figure 1e as compared to 1c. The most tenacious of the properties is the exponential rate of decay. Two rays with different exponential rates of decay are ν-Hölder conjugate, but only for some fixed $\nu < 1$. Applying the widest lens of all, topological conjugacy sees all sinks as the same. In higher dimensions these patterns persist, although the issue of removing rotation becomes more complicated.

To prove the above proposition we begin with an elementary result on Hölder continuous functions.

Lemma 2.2. *The function* $|x|^\nu$, $0 < \nu < 1$, *is Hölder continuous with exponent* ν*; in particular* $||y|^\nu - |x|^\nu| \le |y - x|^\nu$ *for all* $x, y \in \mathbb{R}$.

For each $n \ge 1$ *the function*

$$f(x) = \begin{cases} x(\ln|x|)^n, & x \neq 0 \\ 0, & x = 0 \end{cases}$$

is Hölder continuous in the sense that for each ν, $0 < \nu < 1$, *and each compact set* $C \subset \mathbb{R}$ *there is a constant* K *such that* $|f(y) - f(x)| \le K|y - x|^\nu$ *for all* $x, y \in C$.

Proof. Let $0 < |x| < |y|$ and define $t = |y/x|$, so $1 < t < \infty$. Now

$$\frac{||y|^\nu - |x|^\nu|}{|y - x|^\nu} \le \frac{||y|^\nu - |x|^\nu|}{||y| - |x||^\nu} = \frac{t^\nu - 1}{(t-1)^\nu} \le 1.$$

The last inequality follows from the fact that the rightmost function of t has a limit of zero as $t \to 1$ and a limit of 1 as $t \to \infty$ and has a positive derivative on $(1, \infty)$.

To show the second claim, consider first the function

$$g(u) = \begin{cases} u^\mu (\ln u)^n, & u > 0 \\ 0, & u = 0, \end{cases}$$

where $\mu > 1$. $g(u)$ has a continuous derivative on $u \geq 0$, since

$$g'(0) = \lim_{h \to 0^+} \frac{h^\mu (\ln h)^n}{h} = \lim_{h \to 0^+} (h^{(\mu-1)/n} \ln h)^n = 0$$

and $g'(u) = \mu u^{\mu-1}(\ln u)^n + n u^{\mu-1}(\ln u)^{n-1} \to 0$ as $u \to 0^+$. Thus, on any finite interval $I \subset [0, \infty)$, $|g'(u)| \leq K$ for some constant K and so $|g(v) - g(u)| \leq K|v - u|$ for $u, v \in I$.

Now we want to show f is ν-Hölder for all $0 < \nu < 1$. First let $x, y \geq 0$. Given any such ν, define g as above with $\mu = 1/\nu > 1$. By the properties of ln, we see that $f(x) = \frac{1}{\nu^n} g(x^\nu)$. Therefore, for x, y bounded and nonnegative, $0 \leq x, y \leq M$, there is a constant K such that

$$|f(y) - f(x)| = \frac{1}{\nu^n}|g(y^\nu) - g(x^\nu)| \leq \frac{K}{\nu^n}|y^\nu - x^\nu| \leq \frac{K}{\nu^n}|y - x|^\nu.$$

Since f is an odd function, the same holds for $x, y \leq 0$. If $x < 0 < y$, then

$$\begin{aligned} |f(y) - f(x)| &\leq |f(y) - f(0)| + |f(|x|) - f(0)| \\ &\leq \frac{K}{\nu^n}|y|^\nu + \frac{K}{\nu^n}|x|^\nu \\ &\leq 2\frac{K}{\nu^n}|y + |x||^\nu = 2\frac{K}{\nu^n}|y - x|^\nu \end{aligned}$$

Consequently, $f(x)$ is Hölder continuous. ■

We can now prove Proposition 2.1.

Proof. The matrices A_1 and A_2 are similar, i.e., $P^{-1}A_1 P = A_2$ where

$$P = \begin{bmatrix} 3 & -2 \\ -6 & 6 \end{bmatrix},$$

so the linear change of variables $x_1 = P x_2$ takes the systems $\dot{x}_1 = A_1 x_1$ to the system $\dot{x}_2 = A_2 x_2$. The columns of P are eigenvectors of A_1. The systems (2a) and (2b) are linearly conjugate.

The change of variables

$$\begin{aligned} u_3 &= u_4 \cos\left(\tfrac{1}{4}\ln(u_4^2 + v_4^2)\right) + v_4 \sin\left(\tfrac{1}{4}\ln(u_4^2 + v_4^2)\right), \\ v_3 &= v_4 \cos\left(\tfrac{1}{4}\ln(u_4^2 + v_4^2)\right) - u_4 \sin\left(\tfrac{1}{4}\ln(u_4^2 + v_4^2)\right) \end{aligned} \tag{5}$$

takes the system of equations (2c) to (2d) and thus defines a conjugacy ϕ between the two systems. The inverse is the same except for some changes in sign. It is easier to understand this change of variables in polar coordinates r_k, θ_k. The equations in polar coordinates are

$$\begin{aligned} \dot{r}_3 &= -2r_3 \qquad & \dot{r}_4 &= -2r_4 \\ \dot{\theta}_3 &= 0 \qquad & \dot{\theta}_4 &= -1 \end{aligned}$$

and the change of variables is $r_3 = r_4$, $\theta_3 = \theta_4 - \frac{1}{2}\ln r_4$.

To see that the change of variables in (5) is Lipschitz, first compute the partials away from the origin and note that they are bounded on compact sets. Then from the form of (5) it follows that

$$\|x_3 - 0\| \le |u_3| + |v_3| \le 2|u_4| + 2|v_4| \le 4\|x_4 - 0\|.$$

Thus, systems (2c) and (2d) with phase portraits in Figure 1c and 1d are Lipschitz conjugate.

Now look at systems (2c) and (2e). Consider the change of variables

$$u_5 = u_3, \qquad v_5 = v_3 - \frac{1}{2}u_3 \ln |u_3|, \tag{6}$$

where it is understood that the origin is taken to the origin. The inverse is

$$u_3 = u_5, \qquad v_3 = v_5 + \frac{1}{2}u_5 \ln |u_5|.$$

By Lemma 2.2 these are Hölder continuous and away from the origin they are smooth. If u_3, v_3 satisfies the systems (2c) then

$$\dot{u}_5 = \dot{u}_3 = -2u_3 = -2u_5$$

and

$$\dot{v}_5 = \dot{v}_3 - \frac{1}{2}\dot{u}_3 \ln |u_3| - \frac{1}{2}\dot{u}_3 = -2v_3 + u_3 \ln |u_3| + u_3 = -2v_5 + u_5$$

which shows that (6) takes the system of equations (2c) to (2e). Thus, systems (2c) and (2e), with phase portraits in Figure 1c and 1e, are Hölder conjugate.

Geometrically, (6) takes the straight line trajectory $u_3 = e^{-2t}$, $v_3 = 0$ (the u_3-axis) to the trajectory $u_5 = e^{-2t}$, $v_5 = te^{-2t}$ (the curve $v_5 = -\frac{1}{2}u_3 \ln u_3$), thus breaking the tangencies in Figure 1e.

Again by direct differentiation we see that

$$u_6 = \operatorname{sgn}(u_3)\, |u_3|^{1/2}, \qquad v_6 = \operatorname{sgn}(v_3)\, |v_3|^{1/2} \tag{7}$$

takes the system of equations (2c) to (2f). By Lemma 2.2 it is Hölder continuous of order $\frac{1}{2}$. The inverse transformation is even smoother. Thus, systems (2c) and (2f), with phase portraits in Figure 1c and 1f, are $\frac{1}{2}$-Hölder conjugate. Similarly,

$$u_3 = u_2, \qquad v_3 = \operatorname{sgn}(v_2)\, |v_2|^{2/3}$$

shows that (2c) and (2b) are $\frac{2}{3}$-Hölder conjugate.

All the above transformations are at least homeomorphisms, and so all systems in (2) are topologically conjugate. In fact, they are all ν-Hölder conjugate with exponent $\nu = \frac{1}{2}$. We are left with showing that the conjugacies we have given are the best possible. This will be taken care of by the following proposition, where we restate this half more explicitly and also slightly more generally. ∎

Proposition 2.2. *The statements in Proposition* 2.1 *are the best possible for the smoothness classes we are considering. Specifically,*

- *systems* (2c) *and* (2d)*, with phase portraits in Figures 1c and 1d, are not linearly equivalent,*
- *systems* (2c) *and* (2e)*, with phase portraits in Figures 1c and 1e, are not Lipschitz equivalent,*
- *systems* (2c) *and* (2f)*, with phase portraits in Figures 1c and 1f, are not ν-Hölder conjugate for any $\nu > \frac{1}{2}$. (Likewise, systems* (2c) *and* (2b)*, with phase portraits in Figures 1c and 1b, are not ν-Hölder conjugate for any $\nu > \frac{2}{3}$.)*

Proof. To show these are the smoothest possible, we rely on the geometry apparent in the phase portraits. For the first claim, observe that any equivalence between systems (2c) and (2d) must take the straight line orbits of the ray (1c) onto the curved orbits of the focus (1d). Since a linear transformation takes lines to lines, systems (2c) and (2d) cannot be linearly equivalent.

A critical geometric feature of Lipschitz transformations (i.e., where both ϕ and ϕ^{-1} are Lipschitz) is that they preserves tangencies—see Proposition A.2 in the Appendix. All of the nonzero orbits of (1e) are tangent to the vertical axis, but no two orbits of the ray (1c) are tangent. Thus, systems (2c) and (2e) are not Lipschitz equivalent.

Let $x_6 = f(x_3)$ be a ν-Hölder conjugacy between systems (2c) and (2f) and suppose $\nu > \frac{1}{2}$. Since 0 is the only constant solution under both systems, we must have $f(0) = 0$. Choose any $\bar{x}_3 \neq 0$ and set $\bar{x}_6 = f(\bar{x}_3)$. Let C be a compact set which contains $\{e^{-2t}\bar{x}_3 : t \geq 0\}$ and K the corresponding constant in the definition of Hölder conjugacy. Since f is a homeomorphism, $\bar{x}_6 \neq 0$. f takes the trajectory through $\bar{x}_3$ onto the trajectory through $\bar{x}_6$. Therefore for $t \geq 0$,

$$e^{-t}\bar{x}_6 = f(e^{-2t}\bar{x}_3), \tag{8}$$

$$e^{-t}\|\bar{x}_6\| = \|f(e^{-2t}\bar{x}_3)\| \leq K\|e^{-2t}\bar{x}_3\|^{\nu}, \tag{9}$$

$$\|\bar{x}_6\| \leq e^{(1-2\nu)t}K\|\bar{x}_3\|^{\nu}. \tag{10}$$

However, if $\nu > \frac{1}{2}$, then $(1 - 2\nu) < 0$ and this implies $\bar{x}_6 = 0$, which is a contradiction. Thus, systems (2c) and (2f) cannot be ν-Hölder conjugate for $\nu > \frac{1}{2}$.

A similar argument shows there can be no ν-Hölder conjugacy between systems (2c) and (2b) for $\nu > \frac{2}{3}$. Any conjugacy between these two systems would take the positive u_2-axis onto one of the rays in Figure 1c, which by composition with a rotation in the x_3 plane can be taken to be the positive u_3-axis. Now the same argument as above restricted to the positive u_2 and u_3 axes shows the Hölder exponent cannot be greater than $\frac{2}{3}$. ■

3. THE GENERAL CASE In much of the previous discussion we can simply replace $\mathbb{R}^2$ with $\mathbb{R}^n$. In particular, consider the vector x as an n-vector and A as an $n \times n$ matrix in equation (1). The definitions of trajectory, orbit, conjugate, equivalent, etc. are exactly the same. Of course, the question is what is the generalization of Proposition 2.1. The following theorem reduces the discussion to a few cases.

Theorem 3.1 (The Jordan Canonical Form Theorem). *There exists a real, nonsingular matrix P such that*

$$PAP^{-1} = \mathrm{diag}[J(\lambda_1), J(\lambda_2), \ldots, J(\lambda_k)] \tag{11}$$

where λ_j, $j = 1, \ldots, k$, are eigenvalues of A (possibly repeated) and the $J(\lambda_j)$ are submatrices defined as follows. For a real eigenvalue λ, the submatrix $J(\lambda)$ has the *form*

$$\begin{bmatrix} \lambda & 0 & 0 & \cdots & 0 & 0 \\ 1 & \lambda & 0 & \cdots & 0 & 0 \\ 0 & 1 & \lambda & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda & 0 \\ 0 & 0 & 0 & \cdots & 1 & \lambda \end{bmatrix}, \tag{12}$$

and for a complex eigenvalue $\lambda = \alpha + \beta i$, $J(\lambda) = J(\alpha + \beta i)$ *has the form*

$$\begin{bmatrix} R & 0 & 0 & \cdots & 0 & 0 \\ I & R & 0 & \cdots & 0 & 0 \\ 0 & I & R & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & R & 0 \\ 0 & 0 & 0 & \cdots & I & R \end{bmatrix}, \tag{13}$$

where

$$R = \begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix}, \qquad I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \tag{14}$$

See Hirsch and Smale [**3**] for a discussion of this theorem.

The submatrices $J(\lambda_j)$ are called *Jordan blocks* and PAP^{-1} in (11) is called the *(real) Jordan form of A*. Let $J(\lambda_j)$ be $n_j \times n_j$; then n_j is even if λ_j is complex and $n_1 + n_2 + \cdots + n_k = n$. The n_j can be 1, in which case $J(\lambda_j)$ contributes a single entry on the diagonal. The Jordan form is unique up to the order in which the Jordan blocks appear on the diagonal.

Since A and its real Jordan form are similar, $\dot{x} = Ax$ and $\dot{y} = PAP^{-1}y$ are linearly conjugate. Therefore, we can assume A is already in its real Jordan form. Except for the first, all of the matrices in (3) are in real Jordan form. The matrices (3b), (3c), and (3f) all have two 1×1 real Jordan blocks; (3e) has one 2×2 real Jordan block; and (3d) has one 2×2 complex Jordan block.

Each Jordan block contributes a 1-dimensional eigenspace to the matrix in the case of a real eigenvalue, or a 2-dimensional generalized eigenspace in the case of a complex eigenvalue. The off-diagonal terms of the Jordan block link the variables associated with the block into a cohesive unit. Breaking this structure—removing the off-diagonal blocks—means unlinking these variables.

Since variables associated to one Jordan block are independent of the other variables, we state the theorem for a single Jordan block and we only consider the case where the eigenvalue has nonzero real part. See [**5**] for an indication of what is involved for topological conjugacy in the case when the eigenvalues have zero real part.

Theorem 3.2. *Let* $A = J(\lambda)$, *where* λ *is either real or complex and* $J(\lambda)$ *is as in* (12) *or* (13), *respectively.*

I. *If A has a rotational component, i.e.,* $\mathrm{Im}(\lambda) \neq 0$, *and*

(a) *if A is a trivial* (2×2) *Jordan block, then the rotation can be eliminated by a Lipschitz conjugacy, but not by a linear equivalence.*

(b) *if A is a nontrivial Jordan block, then the rotation can be eliminated by a Hölder conjugacy, but not by a Lipschitz equivalence.*

II. *Breaking the Jordan block structure of A, i.e., removing the off-diagonal blocks in* (12) *or* (13)*, can be achieved with a Hölder conjugacy, but not by a Lipschitz equivalence.*

III. *The rate at which solutions approach the origin can be changed by a ν-Hölder conjugacy only if ν is sufficiently less than* 1. *More explicitly, if $B = J(\mu)$ is a second Jordan block of the same size as A, $\alpha := \mathrm{Re}(\lambda)$, $\beta := \mathrm{Re}(\mu)$, and $0 < \alpha/\beta < 1$, then $\dot{x} = Ax$ and $\dot{y} = By$ are $\frac{\alpha}{\beta}$-Hölder conjugate, but not ν-Hölder conjugate for any $\nu > \alpha/\beta$.*

The properties above are arranged in order of increasing rigidity. Intrinsically the rate of rotation is the least robust property of the system. On its own it can be removed the most easily. However, we will see below that if A has a nontrivial Jordan block structure then this structure becomes entangled with the rotation, and removing the rotation is tantamount to breaking the blocks. Hence, in this case, rotation becomes as rigid as the block structure itself. The most rigid of the three properties is the exponential rate of decay.

Only I(b) and II for complex λ are fundamentally different from what was presented in the previous examples—simply because a nontrivial Jordan block with complex eigenvalue cannot occur in examples smaller than 4×4. This is the most complicated of the arguments and will be presented last. However, it will also provide an interesting illustration of a geometric property preserved by Lipschitz transformations.

The proof of this theorem requires another elementary lemma.

Lemma 3.1. *The composition of two Hölder continuous functions is Hölder continuous. The composition of two Hölder conjugacies is a Hölder conjugacy.*

Proof. If ϕ is a α-Hölder function and ψ is a β-Hölder function, then

$$\|(\phi \circ \psi)(x) - (\phi \circ \psi)(y)\| \leq K_\phi \|\psi(x) - \psi(y)\|^\alpha \leq K_\phi (K_\psi)^\alpha \|x - y\|^{\alpha\beta}$$

Therefore, the composition is only $\alpha\beta$-Hölder. However, if ϕ is α-Hölder for all $\alpha < 1$ and ψ is β-Hölder for all $\beta < 1$, then $\alpha\beta$ can be taken arbitrarily close to 1. Consequently, the composition of two Hölder continuous functions is Hölder continuous (i.e., ν-Hölder for all $\nu < 1$). By extension, the composition of two Hölder conjugacies is a Hölder conjugacy. ■

Since a Lipschitz conjugacy is also a Hölder conjugacy, it follows that the composition of a Hölder conjugacy and a Lipschitz conjugacy, in either order, is a Hölder conjugacy.

For ease of exposition, when convenient we assume the real part of our eigenvalue is negative. As indicated earlier, a positive real part can be dealt with by reversing time. We will work through the theorem first showing the existence of the conjugacies and then that these are optimal. Part I(a) is essentially unchanged from Proposition 2.1. Let $\lambda = \alpha + \beta i$. We want a conjugacy between $\dot{x} = Ax$ and $\dot{y} = By$, where

$$A = \begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix} \qquad \text{and} \qquad B = \begin{bmatrix} \alpha & 0 \\ 0 & \alpha \end{bmatrix}.$$

The generalization of the conjugacy (5) is

$$\begin{aligned} y_1 &= x_1 \cos\left(\frac{\beta}{2\alpha}\ln(x_1^2+x_2^2)\right) - x_2 \sin\left(\frac{\beta}{2\alpha}\ln(x_1^2+x_2^2)\right), \\ y_2 &= x_2 \cos\left(\frac{\beta}{2\alpha}\ln(x_1^2+x_2^2)\right) + x_1 \sin\left(\frac{\beta}{2\alpha}\ln(x_1^2+x_2^2)\right). \end{aligned} \tag{15}$$

Compare (5) to (15) with $\alpha = -2$ and $\beta = 1$. Transformation (15) can be seen to be Lipschitz by the same argument as found in the proof of Proposition 2.1. The inverse transformation has the same form as (15) and so is also Lipschitz.

Part I(b) uses part II, so we take it first. Part II breaks into two cases depending on whether λ is real or complex. Assume first that λ is real. Consider the two systems

$$\dot{x} = Ax, \quad A = J(\lambda); \qquad \dot{y} = By, \quad B = \lambda I \tag{16}$$

where $\lambda \neq 0$. A is the full real Jordan block (12) and B is the block with the off diagonal terms removed. The direct generalization of the transformation in (6) is

$$x_j = \sum_{k=1}^{j} \frac{1}{(j-k)!\,\lambda^{j-k}} y_k (\ln|y_k|)^{j-k}. \tag{17}$$

This is a conjugacy between the two systems and by Lemma 2.2 is Hölder continuous. The inverse transformation can be written recursively as

$$y_1 = x_1; \qquad y_j = x_j - \sum_{k=1}^{j-1} \frac{1}{(j-k)!\,\lambda^{j-k}} y_k (\ln|y_k|)^{j-k}, \; j = 2, \dots, n$$

The right side is Hölder continuous in the y_k ($k < j$). Inductively assume each y_k to be Hölder continuous in the x_i for $1 \leq i \leq k$. This certainly holds for y_1. y_j is then the composition of Hölder continuous functions and by Lemma 3.1 is Hölder continuous in $x_1, \dots, x_j$. Therefore, (17) is a Hölder conjugacy.

Now suppose $\lambda = \alpha + \beta i$, where α and β are nonzero reals. Consider the two systems

$$\dot{x} = Ax, \quad A = J(\alpha + \beta i); \qquad \dot{y} = By \tag{18}$$

where A is the full complex Jordan block in (13) and B is the block-diagonal matrix obtained from A by removing the off-diagonal blocks,

$$B = \begin{bmatrix} R & 0 & \cdots & 0 \\ 0 & R & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & R \end{bmatrix}. \tag{19}$$

Both A and B are in block form with 2×2 blocks, so let $x = (x_1, \dots, x_n)$ and $y = (y_1, \dots, y_n)$, where x_j and y_j are 2-vectors. The equations to be considered are

$$\begin{gathered} \dot{x}_1 = Rx_1; \qquad \dot{x}_j = x_{j-1} + Rx_j, \; j = 2, \dots, n, \\ \dot{y}_j = Ry_j, \; j = 1, \dots, n. \end{gathered} \tag{20}$$

The analog of the transformation (17) is

$$x_j = \sum_{k=1}^{j} \frac{1}{(j-k)!\,\alpha^{j-k}} y_k (\ln \|y_k\|)^{j-k} \tag{21}$$

with recursively defined inverse

$$y_1 = x_1; \qquad y_j = x_j - \sum_{k=1}^{j-1} \frac{1}{(j-k)!\,\lambda^{j-k}} y_k (\ln \|y_k\|)^{j-k},\ j = 2, \dots, n.$$

This is a Hölder conjugacy between the two systems in (20).

To show part I(b), start by applying part II to obtain a Hölder conjugacy between the systems in (18), then apply part I(a) to remove the rotation from each R by a Lipschitz conjugacy, and finally use part II again to replace the off-diagonal blocks via a Hölder conjugacy. By Lemma 3.1, the composition of these three will be a Hölder conjugacy removing the rotation from the nontrivial Jordan block.

The composition of the first two conjugacies in the preceding paragraph shows that for any Jordan block $A = J(\lambda)$ the system $\dot{x} = Ax$ can be reduced to the diagonal system $\dot{x} = \alpha x$, $\alpha = \mathrm{Re}(\lambda)$, by a Hölder conjugacy. To construct a conjugacy between two systems as in part III, first reduce both to their corresponding diagonal systems, and then this becomes a 1-dimensional problem. For two real variables, the systems $\dot{x} = \alpha x$ and $\dot{y} = \beta y$, where $0 < \alpha/\beta < 1$, are conjugated by $x = \mathrm{sgn}(y)|y|^{\alpha/\beta}$, which is $\frac{\alpha}{\beta}$-Hölder. Observe that if there were a conjugacy between the original two systems that was ν-Hölder for $\nu > \alpha/\beta$, the same would be true for the diagonal systems. An argument following along the lines of equations (8–10) would then produce a contradiction, so part III is optimal.

We are left with showing that parts I and II are optimal. As in Proposition 2.2, this will rely on the geometry inherent in the phase portraits. Part I(a) concerns the same two phase portraits, Figures 1c and 1d, as in this earlier proposition—a ray and a focus. As before, there can be no linear equivalence since it would not be possible to take the straight line orbits of the former onto the spirals of the latter. The phase portrait for a one-tangent node, Figure 1e—the phase portrait for a nontrivial Jordan block with real eigenvalue—is also effectively unchanged in higher dimensions. It has a unique one-dimensional eigenspace and all nonzero trajectories are tangent to this eigenspace. When λ is real, part II is concerned with mapping the phase portrait for a one-tangent node onto that of a diagonal system—a higher dimensional ray. Again, a Lipschitz transformation cannot take the tangent trajectories of the first onto the nontangent trajectories of the second, so there is no Lipschitz equivalence.

The remaining two cases, part I(b) and part II for complex λ, are the most interesting because they involve the interaction between the rotation and the off-diagonal blocks in (13). The argument that neither of these can be achieved with a Lipschitz equivalence will again rely on the geometry inherent in the phase portrait—in this case the phase portrait for a nontrivial Jordan block with complex eigenvalue, (13). Therefore, our primary objective will be to understand this phase portrait. The picture is similar for any size Jordan block and we will describe the 4×4 case. Let $\lambda = \alpha + \beta i$, with $\beta \neq 0$, and

$$A = \begin{bmatrix} \alpha & \beta & 0 & 0 \\ -\beta & \alpha & 0 & 0 \\ 1 & 0 & \alpha & \beta \\ 0 & 1 & -\beta & \alpha \end{bmatrix}. \tag{22}$$

In order to describe the phase portrait for $\dot{x} = Ax$, we begin with the following 3×3 system

$$\begin{bmatrix} \dot{u} \\ \dot{v} \\ \dot{w} \end{bmatrix} = \begin{bmatrix} \alpha & 0 & 0 \\ 1 & \alpha & 0 \\ 0 & 0 & \alpha \end{bmatrix} \begin{bmatrix} u \\ v \\ w \end{bmatrix}. \tag{23}$$

This has solutions $u(t) = u_0 e^{\alpha t}$, $v(t) = v_0 e^{\alpha t} + u_0 t e^{\alpha t}$, and $w(t) = w_0 e^{\alpha t}$. Since $w/u \equiv w_0/u_0$, any trajectory starting in a plane $w = ku$ stays in this plane. $w = ku$ describes any plane through the v-axis with the exception of the vw-plane itself, which is a special case. Moreover, since $u(t)$ and $v(t)$ do not depend on w_0, any two trajectories starting on the same vertical line (i.e., differing only in w_0) will be on a common vertical line for every t. That is, the family of vertical lines is invariant—vertical lines flow to vertical lines. Combining these, we can describe the phase portrait for this 3×3 system. In the uv-plane we have the planar system shown in Figure 1e with all trajectories tangent to the v-axis. The vw-plane is a 2-dimensional eigenspace for the matrix and the solutions, $(u, w) = (u_0, w_0)e^{\alpha t}$, are rays as in Figure 1c. See Figure 2a. Any other trajectory sits on a unique plane determined by its initial point and the v-axis. The trajectory will be the curve in this plane that sits over the planar trajectory with the same u_0 and v_0. This is illustrated in Figure 2b. We can see that all trajectories are tangent to the v-axis with the exception of those in the vw-plane (other than the v-axis itself).

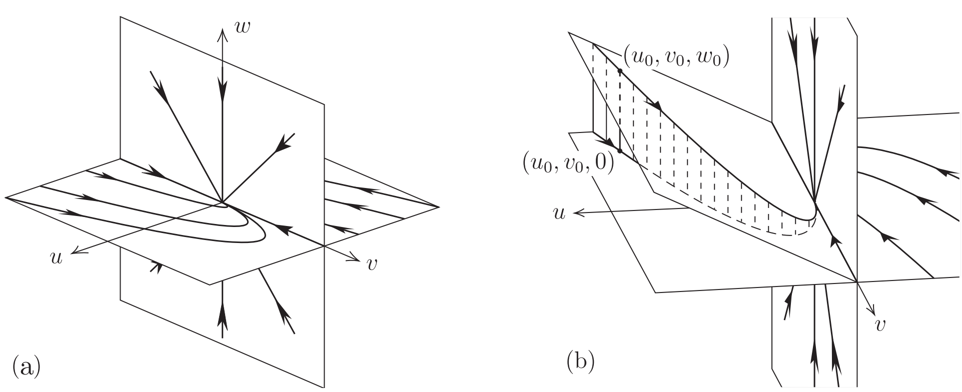


Figure 2. The phase portrait of a 3×3 system.

Now let B be the same as the matrix A except with $\beta = 0$. Let V denote the subspace $x_2 = 0$. V is invariant under $\dot{x} = Bx$ and on V the system reduces (delete the second row and column of B) to the one in (23). Therefore, Figure 2 describes those trajectories of $\dot{x} = Bx$ contained in V. Let $U_\theta = \left[\begin{smallmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{smallmatrix}\right]$ and let $\mathcal{R}_\theta$ be the block-diagonal rotation matrix diag$[U_\theta, U_\theta]$. $\mathcal{R}_\theta$ and B commute, so $\dot{x} = Bx$ is invariant under rotation by $\mathcal{R}_\theta$: if $y = \mathcal{R}_\theta x$, then $\dot{y} = \mathcal{R}_\theta \dot{x} = \mathcal{R}_\theta B \mathcal{R}_\theta^{-1} y = By$. The two satisfy the same differential equation, so they have the same trajectories. This implies that for any θ, $\mathcal{R}_\theta$ takes the trajectories of $\dot{x} = Bx$ (as a group) onto themselves. $\mathcal{R}_\theta$ acts by separately rotating the x_1x_2- and x_3x_4-planes by angle θ. Consequently, given any initial condition $x(0)$, there will be some θ for which $\mathcal{R}_\theta x(0) \in V$ (i.e., $x_2 = 0$). This says that for every $\theta \in S^1$, $\mathcal{R}_\theta V$ is another invariant subspace in which the orbits look like Figure 2 and $\mathbb{R}^4 = \cup_{\theta \in S^1} \mathcal{R}_\theta V$. The two-dimensional eigenspace $(x_1, x_2) = (0, 0)$

is taken to itself under these rotations, so it is common to all of the $\mathcal{R}_\theta V$ and corresponds in each to the vw-plane. From this we can conclude that every nonzero orbit is tangent to (or is) a ray in the eigenspace. Moreover, as θ varies from 0 to 2π, the vector that corresponds to the positive v-axis ($\mathcal{R}_\theta \hat{e}_3$, for standard basis vectors $\{\hat{e}_1, \hat{e}_2, \hat{e}_3, \hat{e}_4\}$) makes one complete revolution in the eigenspace. Therefore, if we let S be the set of nonzero orbits tangent to the positive x_3-axis (the positive v-axis for $\theta = 0$), including the positive x_3-axis itself, then on the one hand, $\cup_{\theta \in S^1} \mathcal{R}_\theta S$ must include all nonzero orbits because it includes all orbits tangent to any eigendirection, and on the other hand, the $\mathcal{R}_\theta S$ must be disjoint since θ defines the direction along which the orbits in $\mathcal{R}_\theta S$ approach the origin. Consequently, the set of all nonzero orbits of $\dot{x} = Bx$ decomposes as the disjoint union of a 1-parameter family of rigid rotations of S, $\{\mathcal{R}_\theta S : \theta \in S^1\}$. Each $\mathcal{R}_\theta S$ looks like roughly half of the orbits in Figure 2 (the rest being the trajectories tangent to the negative v-axis plus the remaining rays in the vw-plane). If we define an equivalence relation on the set of nonzero orbits by defining two orbits to be equivalent when they are tangent to each other at the origin, then the sets $\mathcal{R}_\theta S$ are the equivalence classes. These are the maximal sets of mutually tangent orbits. The equivalence classes are "identical" in the sense that one can be obtained from another by a rigid rotation preserving the full orbit structure. Each equivalence class contains infinitely many orbits, including exactly one which is a ray in the eigenspace, setting up a one-to-one correspondence between the equivalence classes and the rays to which they are tangent, or equivalence classes and angles in S^1.

We now arrive at the system $\dot{x} = Ax$. Let $C := A - B$; C is the matrix consisting of only the β terms. B and C commute, so the solutions of $\dot{x} = Ax$ are $x(t) = e^{At}x(0) = e^{Ct}e^{Bt}x(0)$. Since $e^{Ct} = \mathcal{R}_{\beta t}$, we have $x(t) = \mathcal{R}_{\beta t}e^{Bt}x(0)$. $e^{Bt}x(0)$ is the corresponding solution to $\dot{x} = Bx$. Therefore, if we strip off the rotational component by looking at the system in the rotating coordinate system, $x \mapsto \mathcal{R}_{\beta t}x$, the solutions to $\dot{x} = Ax$ look like those of $\dot{x} = Bx$. Specifically, if $x(0) \in \mathcal{R}_\theta S$, so $\mathcal{R}_\theta^{-1}x(0) = (u_0, 0, v_0, w_0)$, and we define the (partial) rotating frame $\{\hat{\mathbf{u}}, \hat{\mathbf{v}}, \hat{\mathbf{w}}\} := \mathcal{R}_{\beta t+\theta}\{\hat{e}_1, \hat{e}_3, \hat{e}_4\}$, then

$$x(t) = u_0 e^{\alpha t}\hat{\mathbf{u}} + (v_0 + u_0 t)e^{\alpha t}\hat{\mathbf{v}} + w_0 e^{\alpha t}\hat{\mathbf{w}}. \tag{24}$$

As time evolves the sets $\{\mathcal{R}_\theta S\}$ flow one into another cyclically, and within a moving representative the flow looks like that in Figure 2.

Because of the rotation, orbits are no longer tangent to a ray, but are still tangent to the eigenspace. If we look at the projection of orbits into the eigenspace, the uniform rotation means that all orbits will wrap around the origin. However, we claim that orbits do not all wrap with equal density and this will be the obstruction to finding the Lipschitz equivalences. Since the rate of rotation is the same for all orbits, how tightly wrapped an orbit is depends upon the rate at which it approaches the origin. Therefore, we want to show that different trajectories can approach the origin at dramatically different rates.

The rate at which a trajectory approaches the origin is not affected by the rotation, $\mathcal{R}_{\beta t}$. Therefore, it will suffice to look at the solutions of $\dot{x} = Bx$. The behavior of $\dot{x} = Bx$ is identical on each of the sets $\mathcal{R}_\theta S$ and on each is described by (23), as illustrated in Figure 2. This reduces the problem to looking at the rates at which trajectories tangent to the positive v-axis approach the origin. Since the trajectories are tangent to the v-axis, asymptotically the distance from the origin is given by the v-coordinate. However, for our purposes what will be important is the distance as measured in the eigenspace. That is, if $\pi(u, v, w) = (v, w)$ is the usual projection and $x(t)$ is any trajectory, then we want to monitor the distance $\|\pi \circ x(t)\|$. Alternatively, if $C(r)$ denotes the cylinder of radius r parallel to the u-axis, $v^2 + w^2 = r^2$, then we are interested in the rate at which $x(t)$ crosses these cylinders. In spite of the fact that all of these trajectories are tangent, we claim that the trajectory on the v-axis approaches the origin faster in this sense than any other. In fact, any other trajectory can be started as close to the origin as we wish, and the trajectory on the v-axis will catch up to it and pass it in a finite amount of time. More important for our purposes, if these two trajectories are both started on the same cylinder $C(r_0)$ and we look at the difference in the time it takes the two trajectories to reach a smaller cylinder, $C(r)$, then this difference can be made as large as we wish by taking r sufficiently small. We state this as a lemma. (Here stated assuming $\alpha < 0$.)

Lemma 3.2. *Let $\mathcal{O}$ be any orbit of* (23) *tangent to the positive v-axis that is not the axis itself. Suppose $r_0 > 0$ is sufficiently small that $\mathcal{O}$ crosses each cylinder $C(r)$, $r \le r_0$, only once (i.e., close enough that $\mathcal{O}$ is already nearly parallel to the v-axis). Let $\bar{x}(t)$ be the trajectory on the positive v-axis with $\bar{x}(0) \in C(r_0)$ ($\bar{x}(0) = (0, r_0, 0)$). Given any r_1 with $0 < r_1 < r_0$, let $x(t)$ be the trajectory on $\mathcal{O}$ with $x(0) \in C(r_1)$ ($\|\pi \circ x(t)\| = r_1$). There exists a $T > 0$ such that for all $t > T$, $\|\bar{x}(t)\| < \|\pi \circ x(t)\|$.*

Alternatively, suppose $\bar{x}(t)$ is as before and $x(t)$ is the trajectory on $\mathcal{O}$ with $x(0) \in C(r_0)$ ($\|\pi \circ x(t)\| = r_0$). For any r with $0 < r \le r_0$, let $\tau(r)$ and $\bar{\tau}(r)$ be the times at which $x(t)$ and $\bar{x}(t)$, respectively, reach the cylinder $C(r)$. Then $\tau - \bar{\tau} \to +\infty$ as $r \to 0^+$.

Proof. If $x(0) =: (u_0, v_0, w_0)$, then our two trajectories are

$$x(t) = (u_0 e^{\alpha t}, (u_0 t + v_0)e^{\alpha t}, w_0 e^{\alpha t}) \quad \text{and} \quad \bar{x}(t) = (0, r_0 e^{\alpha t}, 0).$$

To prove the first claim, observe that

$$(\|\pi \circ x(t)\| / \|\bar{x}(t)\|)^2 = ((u_0 t + v_0)^2 + w_0^2)/r_0^2 \to +\infty$$

for any $x(0)$ provided $u_0 \ne 0$, which is true for any orbit tangent to the v-axis that is not the v-axis itself.

To prove the second part, for brevity let $\tau := \tau(r)$, $\bar{\tau} := \bar{\tau}(r)$ and $x := x(\tau)$, $\bar{x} := \bar{x}(\bar{\tau})$. See Figure 3. Let $\hat{x} := (u_0\tau + v_0, w_0)$, so that $\pi(x) = \hat{x}e^{\alpha\tau}$. Since $x(t)$ cannot reach the origin in finite time, $\tau \to +\infty$ as $r \to 0^+$, which implies $\|\hat{x}\| \to +\infty$. Now $r = \|\pi(x)\| = \|\hat{x}\| e^{\alpha(\tau-\bar{\tau})} e^{\alpha\bar{\tau}}$ and also $r = \|\bar{x}\| = r_0 e^{\alpha\bar{\tau}}$. Therefore, $\|\hat{x}\| e^{\alpha(\tau-\bar{\tau})} = r_0$. Since $\|\hat{x}\| \to +\infty$ and r_0 is constant, we must have $\tau - \bar{\tau} \to +\infty$ (since we are assuming $\alpha < 0$). ∎

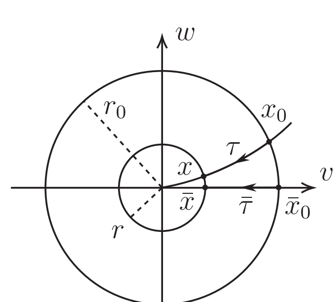


Figure 3. Time difference between trajectories.

Observe that since all trajectories in the eigenspace (vw-plane) approach the origin at the same rate, we would have had the same conclusions if $\bar{x}(t)$ had been a trajectory on any of the rays in the eigenspace, not just the ray to which $\mathcal{O}$ was tangent. Now let $x(t)$ be any trajectory of $\dot{x} = Ax$ not in the eigenspace and let $\bar{x}(t)$ be any nonzero trajectory of $\dot{x} = Ax$ that is in the eigenspace such that $x(0), \bar{x}(0) \in C(r_0)$, where $C(r) \subset \mathbb{R}^4$ is the cylinder $x_3^2 + x_4^2 = r^2$ and r_0 is as in the preceding lemma. Since the rotation does not affect the rate at which trajectories approach the origin, we have from the above lemma and observation that if τ and $\bar{\tau}$ are defined as before, then $\tau - \bar{\tau} \to +\infty$ as $r \to 0^+$. Let $\theta(t)$ be the angle in polar coordinates for the projection of $x(t)$ into the eigenspace and $\bar{\theta}(t)$ the angle for $\bar{x}(t)$. In time t, these angles change by βt. Therefore, when traversing the region from $C(r_0)$ to $C(r)$, the difference in these two angles changes by $\beta(\tau - \bar{\tau}) \to +\infty$ as $r \to 0^+$. Consequently, there exists a sequence $r_n \to 0^+$ such that $\theta(\tau(r_n)) - \bar{\theta}(\bar{\tau}(r_n)) = n\pi$. That is, the two trajectories are alternately at the same point and at antipodal points on the circles of radius r_n on projecting $x(t)$ into the eigenspace. In particular, $x(t)$ repeatedly "laps" $\bar{x}(t)$.

In the eigenspace, $\dot{x} = Ax$ reduces to the two-dimensional focus illustrated in Figure 1d, so apply the Lipschitz transformation from (15) to the x_3, x_4-coordinates to eliminate the rotation in the eigenspace. Let $y(t)$ and $\bar{y}(t)$ denote the images of $x(t)$ and $\bar{x}(t)$, respectively, in the transformed system. $\bar{y}(t)$ is now on a ray, but the transformation acts by rigidly rotating a circle of radius r by an angle depending on r, so the angular separation between the two trajectories where they cross the circle of radius r, $\theta(\tau(r)) - \bar{\theta}(\bar{\tau}(r))$, is unchanged under the transformation and $y(t)$ continues to lap $\bar{y}(t)$. Consequently, the underlying orbit for $y(t)$ cannot be tangent to the ray on which $\bar{y}(t)$ resides. Since $y(t)$ represented any trajectory not in the eigenspace and $\bar{x}(t)$ any trajectory in it, we can conclude that after the Lipschitz transformation the rays in the eigenspace are orbits (with well-defined tangent direction at the origin), but they have no other orbits tangent to them.

We can now make our final two arguments for the optimality of part I(b) and part II for complex λ of Theorem 3.2. As in the two-dimensional case, these will rely on the presence or absence of tangencies among orbits. Note that anything Lipschitz equivalent to $\dot{x} = Ax$ must also be Lipschitz equivalent to the above transformed system and vice versa. Therefore, it will suffice to look at what is Lipschitz equivalent to the transformed system.

First, the transformed system cannot be Lipschitz equivalent to $\dot{x} = Bx$ (part I(b)—removing the rotation from $\dot{x} = Ax$). If there were a Lipschitz equivalence, any ray in the transformed system would need to be the image of some orbit of $\dot{x} = Bx$. Every orbit of $\dot{x} = Bx$ has infinitely many other orbits tangent to it and, by Proposition A.2, these orbits would need to be mapped to orbits of the transformed system tangent to the ray. However, we just observed that there are no orbits tangent to the rays, so this is not possible.

To show part II, let $u_n := y(\tau(r_{2n}))$ and $v_n := \bar{y}(\bar{\tau}(r_{2n}))$. Since $\theta(\tau(r_{2n})) - \bar{\theta}(\bar{\tau}(r_{2n}))$ is an even multiple of π, v_n is the projection of u_n into the eigenspace. $x(t)$, and so $y(t)$, comes in tangent to the eigenspace, so $\lim_{n\to\infty} \|u_n - v_n\|/\|v_n - 0\| = 0$. By Lemma A.1 this property must be preserved under any Lipschitz equivalence. Suppose part II were not optimal and there were a Lipschitz equivalence that could remove the off-diagonal blocks from A. Once the off-diagonal blocks are removed the system reduces to a collection of independent 2×2 foci, which by part I(a) can be diagonalized by a Lipschitz equivalence. Stringing these together we end up with a Lipschitz equivalence between the transformed system above and the system $\dot{x} = \alpha I x$, where all the orbits are rays. Under this equivalence $\{u_n\}$ and $\{v_n\}$ would be taken to sequences $\{u_n'\}$ and $\{v_n'\}$ on distinct rays (orbits) separated by some angle $\phi > 0$. By the law of sines, $\|u_n' - v_n'\|/\|v_n' - 0\| \geq \sin\phi > 0$ and so $\lim_{n\to\infty} \|u_n' - v_n'\|/\|v_n' - 0\| \neq 0$, which is a contradiction.

APPENDIX A. TANGENCIES. In this section we show that a Lipschitz map with Lipschitz inverse preserves tangencies provided the image of at least one of the curves has a tangent direction at the point of interest. Let $f : U \to \mathbb{R}^n$ be a homeomorphism of an open subset of $\mathbb{R}^n$ onto its image where both f and f^{-1} are Lipschitz. This means there exists a $K > 0$ such that for all $x, y \in U$

$$\frac{1}{K}\|x - y\| \leq \|f(x) - f(y)\| \leq K\|x - y\|.$$

The behavior of a Lipschitz transformation is explained by the following lemma.

Lemma A.1. *If $(x_n)_{n\geq 1}, (y_n)_{n\geq 1} \subset U$ are sequences converging to a point $p \in U$ such that*

$$\lim_{n\to\infty} \frac{\|x_n - y_n\|}{\|x_n - p\|} = 0,$$

then the sequences $(f(x_n))_{n\geq 1}, (f(y_n))_{n\geq 1}$ converge to $q := f(p)$ in the same manner.

Proof.

$$\frac{\|f(x_n) - f(y_n)\|}{\|f(x_n) - f(p)\|} \leq \frac{K\|x_n - y_n\|}{\frac{1}{K}\|x_n - p\|} = K^2\frac{\|x_n - y_n\|}{\|x_n - p\|}. \qquad \blacksquare$$

Appearances to the contrary, there is no special role played here by $(x_n)_{n\geq 1}$. Since

$$\|x_n - p\| - \|x_n - y_n\| \leq \|y_n - p\| \leq \|x_n - p\| + \|x_n - y_n\|,$$

if $\|x_n - y_n\|/\|x_n - p\| \to 0$, then $\|y_n - p\|/\|x_n - p\| \to 1$, so $\|x_n - y_n\|/\|y_n - p\| \to 0$, and the sequences are interchangeable. Consider the triangle formed by the three points x_n, y_n, and p. Let θ_n be the angle at vertex p; $\theta_n = \angle y_n p x_n$. Since $\|y_n - p\|/\|x_n - p\| \to 1$, these approach isosceles triangles. $\|x_n - y_n\|/\|x_n - p\| \to 0$ implies $\theta_n \to 0$, so the triangles are collapsing. The lemma says this relationship among the points is preserved under a Lipschitz transformation. Whether or not two curves are tangent at a common endpoint, p, can be characterized by sequences of this type. Here we need a definition of "tangent" that does not depend upon a smooth parametrization, since our map is only Lipschitz.

Definition A.1. Let p be an endpoint of a curve $C \subset \mathbb{R}^n$ and let $\hat{u} \in \mathbb{R}^n$ be a unit vector. We say that C is tangent to the direction $\hat{u}$ at p if for every $\epsilon > 0$ there exists $\delta > 0$ such that for all $x \in C$ with $0 < \|x - p\| < \delta$, we have $\|\hat{v} - \hat{u}\| < \epsilon$, where $\hat{v} := (x - p)/\|x - p\|$ is the unit vector in the direction $x - p$. Two simple curves $C_1, C_2 \subset \mathbb{R}^n$ are tangent to each other at a common endpoint p if both are tangent to the same direction at p.

This definition says that the portion of the curve in a sufficiently small neighborhood of p can be contained in a cone about $\hat{u}$ of arbitrarily small angle. Given a curve with a smooth parameterization, one can see from the definition of the derivative and the continuity of $\|\cdot\|$ that the curve is indeed tangent (according to the above definition) to its unit tangent vector (defined via the derivative). Consequently, two curves that are tangent at a common endpoint under the usual definition are also tangent by the above definition.

Proposition A.1. *Let $C_1, C_2 \subset \mathbb{R}^n$ be continuous curves intersecting at a common endpoint, p. If C_1 and C_2 are tangent at p, then given any sequence $(y_n)_{n\geq 1} \subset C_2 \setminus \{p\}$ with $y_n \to p$, there exists a sequence $(x_n)_{n\geq 1} \subset C_1$, $x_n \to p$, with the property that* $\lim_{n\to\infty} \frac{\|x_n - y_n\|}{\|x_n - p\|} = 0$.

Conversely, if C_1 is tangent to the direction $\hat{u}$ and C_2 is not, either because it is tangent to some other direction or is not tangent to any direction at all, then there exists a sequence $(y_n)_{n\geq 1} \subset C_2$ with $y_n \to p$ and $\epsilon_0 > 0$ such that for any sequence $(x_n)_{n\geq 1} \subset C_1$ with $x_n \to p$ we have $\liminf_{n\to\infty} \frac{\|x_n - y_n\|}{\|x_n - p\|} \geq \epsilon_0$.

Proof. First suppose C_1 and C_2 are both tangent to $\hat{u}$ at p and $(y_n)_{n\geq 1} \subset C_2 \setminus \{p\}$ is any sequence converging to p. For each n, choose $x_n \in C_1$ so that $\|x_n - p\| = \|y_n - p\| =: d$, which we can do because C_1 is connected. Clearly $x_n \to p$. We then have

$$\frac{\|x_n - y_n\|}{\|x_n - p\|} = \left\| \frac{x_n - p}{d} - \frac{y_n - p}{d} \right\| \leq \left\| \frac{x_n - p}{\|x_n - p\|} - \hat{u} \right\| + \left\| \frac{y_n - p}{\|y_n - p\|} - \hat{u} \right\|$$

and by assumption both terms on the right have limit 0 as $n \to \infty$.

Now suppose only C_1 is tangent to $\hat{u}$ at p. If C_2 is not tangent to $\hat{u}$, then there must exist an $\epsilon > 0$ such that for any $\delta_n := 1/n$, there exist points $y_n \in C_2$ with $0 < \|y_n - p\| < \delta_n$ and for which $\left\| \frac{y_n - p}{\|y_n - p\|} - \hat{u} \right\| \geq \epsilon$. Since these are both unit vectors, the minimum separation corresponds to a minimum angle, θ_0, given by $\epsilon^2 = 2(1 - \cos\theta_0)$ from the law of cosines. Bounding $\|x_n - y_n\|/\|x_n - p\|$ from below we have

$$\frac{\|x_n - y_n\|}{\|x_n - p\|} = \left\| \frac{x_n - p}{\|x_n - p\|} - \frac{y_n - p}{\|x_n - p\|} \right\| \geq \left\| \frac{y_n - p}{\|x_n - p\|} - \hat{u} \right\| - \left\| \frac{x_n - p}{\|x_n - p\|} - \hat{u} \right\|.$$

Since C_1 is tangent to $\hat{u}$, the last term has limit 0. The preceding term is the distance from the point $\hat{u}$ to some point on the line through $y_n - p$. The minimum distance to any point on this line is $\sin\theta \geq \sin\theta_0$, where θ is the angle between $\hat{u}$ and $y_n - p$. Consequently, $\liminf_{n\to\infty}(\|x_n - y_n\|/\|x_n - p\|) \geq \sin\theta_0 =: \epsilon_0$. ■

Combining the preceding results we get our objective:

Proposition A.2. *Let $C_1, C_2 \subset \mathbb{R}^n$ be continuous curves intersecting at a common endpoint, p. Suppose these are mapped under a homeomorphism f to curves $\widehat{C}_1, \widehat{C}_2 \subset \mathbb{R}^n$, where both f and f^{-1} are Lipschitz. If C_1 and C_2 are tangent at p, and $\widehat{C}_1$ is tangent to some direction at $f(p)$, then $\widehat{C}_1$ and $\widehat{C}_2$ are tangent at $f(p)$.*

Proof. If $\widehat{C}_1$ and $\widehat{C}_2$ are not tangent at $f(p)$, then we can pick $(\hat{y}_n)_{n\geq 1} \subset \widehat{C}_2 \setminus \{f(p)\}$, as in Proposition A.1, such that for any $(\hat{x}_n)_{n\geq 1} \subset \widehat{C}_1$, we have $\|\hat{x}_n - \hat{y}_n\|/\|\hat{x}_n - f(p)\| \nrightarrow 0$. Let $y_n := f^{-1}(\hat{y}_n)$. Then $(y_n)_{n\geq 1} \subset C_2 \setminus \{p\}$ and $y_n \to p$. Since C_1 and C_2 are tangent at p, there exists a sequence $(x_n)_{n\geq 1} \subset C_1$ such that $\|x_n - y_n\|/\|x_n - p\| \to 0$. By the preceding lemma, if $\hat{x}_n := f(x_n)$, then $\|\hat{x}_n - \hat{y}_n\|/\|\hat{x}_n - f(p)\| \to 0$, which contradicts the choice of the sequence $\hat{y}_n$. Consequently, $\widehat{C}_1$ and $\widehat{C}_2$ must be tangent at $f(p)$. ■

REFERENCES

1. P. Blanchard, R. L. Devaney, and G. R. Hall, *Differential Equations*, 2nd ed., Brooks/Cole, New York, 2003.
2. W. E. Boyce and R. C. DiPrima, *Elementary Differential Equations*, 7th ed., John Wiley, New York, 2001.
3. M. W. Hirsch and S. Smale, *Differential Equations, Dynamical Systems, and Linear Algebra*, Academic Press, New York, 1974.
4. J. H. Hubbard and B. H. West, *Differential Equations: A Dynamical Systems Approach*, Springer-Verlag, New York, 1990.
5. N. H. Kuiper and J. W. Robbin, Topologial classification of linear endomorphisms, *Invent. Math.* **19** (1973) 83–106.
6. R. C. Robinson, *An Introduction to Dynamical Systems: Continuous and Discrete*, Pearson Education, Upper Saddle River, NJ, 2004.

PATRICK MCSWIGGEN received his A.B. in Mathematics in 1981 and his Ph.D. in Mathematics under the supervision of Charles Pugh in 1990 from the University of California at Berkeley. He is currently Associate Professor of Mathematics and Assistant Head in the Department of Mathematical Sciences at the University of Cincinnati. In his spare time he enjoys hiking and other mostly dry outdoor activities and traveling with his wife Tara and children Colin, Kevin, and Brian.
Department of Mathematical Sciences, University of Cincinnati, Cincinnati, OH 45221
pat.mcswiggen@uc.edu

KENNETH R. MEYER received his B.E.P. in Engineering Physics from Cornell University in 1960 and his Ph.D. in Mathematics from the University of Cincinnati in 1964. He is Charles Phelps Taft Professor Emeritus at the University of Cincinnati.
Department of Mathematical Sciences, University of Cincinnati, Cincinnati, OH 45221
ken.meyer@uc.edu

Mathematics Is . . .

"Mathematics is the loom upon which God weaves the fabric of the universe."

Clifford A. Pickover, *The Loom of God: Mathematical Tapestries at the Edge of Time*, Plenum Press, New York, 1997, p. 16.

—Submitted by Carl C. Gaither, Killeen, TX