A geometric approach to the Global Attractor Conjecture

Manoj Gopalkrishnan, Ezra Miller, and Anne Shiu

16 December 2013

Abstract

This paper introduces the class of *strongly endotactic networks*, a subclass of the endotactic networks introduced by G. Craciun, F. Nazarov, and C. Pantea. The main result states that the global attractor conjecture holds for complex-balanced systems that are strongly endotactic: every trajectory with positive initial condition converges to the unique positive equilibrium allowed by conservation laws. This extends a recent result by D. F. Anderson for systems where the reaction diagram has only one linkage class (connected component). The results here are proved using differential inclusions, a setting that includes power-law systems. The key ideas include a perspective on reaction kinetics in terms of combinatorial geometry of reaction diagrams, a projection argument that enables analysis of a given system in terms of systems with lower dimension, and an extension of Birch's theorem, a well-known result about intersections of affine subspaces with manifolds parameterized by monomials.

Contents

1 Introduction

The general study of reaction systems with mass-action kinetics goes back at least to the pioneering work of Feinberg, Horn, and Jackson in the 1970s [19, 27]. The class of systems under consideration in this paper includes such mass-action systems, as well as power-law systems and S-systems from biochemical systems theory [43].

Quite apart from the context of biochemistry, such systems appear to be objects of intrinsic interest for engineers and mathematicians. For instance, essentially identical models have been investigated in genetic algorithms [41] and population dynamics [31]. Models with similar mathematical structure have been studied in distributed systems [40] and algebraic statistics [37]. Connections between these systems and binomial algebra and toric geometry have been stressed by several authors [1, 13, 14, 20, 23, 33, 45].

The Global Attractor Conjecture (GAC) is the focus of this paper. Given a reaction system, conservation laws induce a foliation of the concentration space (which is a positive orthant) by polyhedra that are forward-invariant with respect to the dynamics. It is well-known that when the dynamics of a reaction system admit a "pseudo-Helmholtz free energy function" as a strict Lyapunov function, then each forward-invariant polyhedron contains a special equilibrium point, sometimes called the "Birch point" due to a connection to Birch's theorem [27]. The GAC, which has resisted attempts at proof for four decades [25], asserts that every trajectory asymptotically approaches the unique Birch point in its forward-invariant polyhedron. A survey of literature relevant to the GAC appears in Section 4.

The GAC is usually stated for complex-balanced systems, which are generalizations of the more well-known detailed-balanced systems. Complex-balanced systems are known to

admit pseudo-Helmholtz free energy functions as Lyapunov functions. Specializing the GAC to the case of networks in which every reaction has the form $A \to B$ (that is, every complex is a species) yields the well-known Ergodic Theorem for continuous-time, discrete-space, autonomous Markov chains [38]. From this perspective, the GAC may be viewed as an ergodicity conjecture for chemical reaction networks evolving under mass-action kinetics. Indeed, reaction systems theory may be viewed as a nonlinear generalization of the theory of continuous-time discrete-space autonomous Markov chains.

A key idea in our work is that the combinatorial geometry of a reaction network in the space of (chemical) complexes—the reactants and products of the reactions—informs the dynamics in concentration space. This connection was anticipated to a certain extent by the "extended permanence conjecture" (Conjecture 4.5) of Craciun, Nazarov, and Pantea, which implies the GAC [15]. Their conjecture captures the intuitively appealing idea that if a reaction network "points inwards", in the sense of being endotactic (see Definition 3.14), then the corresponding dynamics in concentration space must also roughly "point inwards", in the sense of being permanent (see Definition 4.1).

We develop the correspondence between geometry of a network in the space of complexes and dynamics in concentration space by analyzing the contributions of each reaction to the dynamics along "toric jets" (Definition 6.2.3). In doing so, we positively resolve the extended permanence conjecture for a subclass of endotactic networks called "strongly endotactic" networks (Definition 3.14), our first main result.

Theorem 1.1. *Every strongly endotactic reaction network is permanent.*

As stated above, the extended permanence conjecture implies the GAC.

Theorem 1.2. *The GAC holds for strongly endotactic complex-balanced reaction systems.*

The fact that weakly reversible networks with one linkage class (i.e., with strongly connected reaction graph) are strongly endotactic (Corollary 3.20) yields an easy consequence.

Theorem 1.3. *Every weakly reversible reaction network with exactly one linkage class is permanent.*

Theorem 1.3 strengthens a result in which Anderson proved persistence of weakly reversible systems with exactly one linkage class under the assumption that all trajectories are bounded [4]. Our approach was significantly influenced by distillation of ideas from Anderson's result, combined with insights gained from [15].

The class of dynamical systems considered here includes those known as generalized mass-action systems, or power-law systems, studied in biochemical systems theory [43]. More precisely, we prove our results here in the more general setting of mass-action differential inclusions, which we introduced in earlier work [24]. The crucial result from [24] applied here (in Section 8) concerns families of differential inclusions that are closed under certain projections, allowing us to analyze a system in terms of systems of lower dimension.

In the course of this work (Section 5), we prove two extensions of Birch's theorem, a well-known result in reaction network theory [27] and algebraic statistics [11] concerning intersections of certain affine spaces and manifolds parametrized by monomials (Theorems 5.12 and 5.14). We show that Birch's theorem remains true under slight perturbation (Theorem 5.17), and at infinity in suitable compactifications (Theorem 5.16).

To connect the combinatorial (polyhedral) geometry of a reaction diagram to the asymptotics of its dynamics (Section 6), we develop jets and toric jets (Definition 6.2), in parallel with the notion of "jet frame" from work by Miller and Pak [34] on unfolding convex polyhedra. Jets capture motion the bulk of which proceeds in a fixed main direction, but is perturbed to first order in an orthogonal direction, to second order in a third direction, and so on. Jets coherently tease apart the contributions of various reactions to the gradient of pseudo-Helmholtz free energy along infinite trajectories (Proposition 6.26).

Birch's Theorem and jets come together (in Section 7) to show that for strongly endotactic networks, within each stoichiometric compatibility class there exists a compact set outside of which the pseudo-Helmholtz free energy function $\sum_{i\in\mathcal{S}}(x_i \log x_i - x_i)$ decreases along trajectories (Theorem 7.5). Our main results on persistence and permanence, described earlier in this Introduction, follow in Section 8.

Examples in Section 9 illustrate our results and explain the limitations of our approach. A further extended example in Section 2 serves to introduce the main ideas, after which we give precise, general definitions concerning reaction networks and their accompanying mass-action differential inclusions (Section 3). Various conjectures related to the GAC, along with known partial results, are collected in Section 4.

2 An illustrative example

Consider the following reaction network with two species X and Y and three reactions:

$$2X \xrightarrow{k_1} X \qquad 0 \xrightarrow{k_2} Y \qquad 2Y \xrightarrow{k_3} X + Y, \tag{1}$$

where $k_1, k_2, k_3 \in \mathbb{R}_{>0}$ denote the reaction rate constants. Network (1) is obtained by reversing all the reactions in the well-known Lotka–Volterra reaction network. Letting $x(t)$ and $y(t)$ denote concentrations of X and Y, respectively, at time t, network (1) defines the following system of ordinary differential equations arising from mass-action kinetics:

$$\begin{pmatrix}\dot{x}\\ \dot{y}\end{pmatrix} = k_1 x^2 \begin{pmatrix}-1\\ 0\end{pmatrix} + k_2 \begin{pmatrix}0\\ 1\end{pmatrix} + k_3 y^2 \begin{pmatrix}1\\ -1\end{pmatrix}. \tag{2}$$

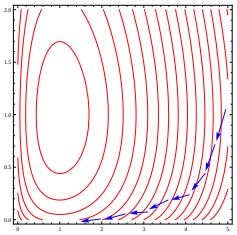
In this section, we introduce the main ideas of our work by explaining how to prove that network (1) taken with mass-action kinetics is *permanent*: there exists a compact set $K \subseteq \mathbb{R}^2_{>0}$ such that every trajectory of the dynamical system (2) in $\mathbb{R}^2_{>0}$ eventually remains in K. We begin with the following assertion.

Claim 2.1. There exists a compact set $K \subseteq \mathbb{R}^2_{>0}$ (that depends on k_1, k_2, k_3) such that outside K, the function

$$g(x, y) = x \log x - x + y \log y - y$$

on $\mathbb{R}^2_{>0}$ is strictly decreasing along trajectories $\big(x(t), y(t)\big)$ of (2) except at an equilibrium.

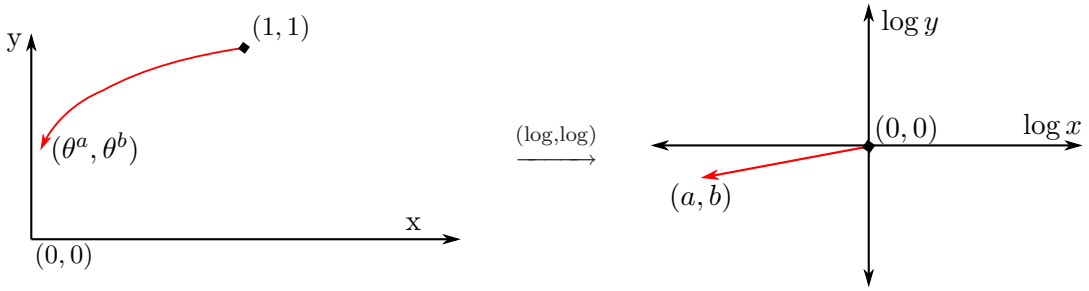
By itself, Claim 2.1 does not imply that the dynamical system (2) is permanent. For instance, consider the figure



in which level sets of g are indicated in red, with interior level sets correspond to lower values of g. The function g decreases along the trajectory indicated in blue, but this trajectory is not even persistent—that is, at least one coordinate (in this case, the y-coordinate) approaches zero.

Nevertheless, Claim 2.1 is enough to establish that the origin $(0, 0)$ is repelling (Definition 4.1.2) and that all trajectories of (2) are bounded. These two properties, along with good behavior under projection, suffice to obtain our desired permanence result. This follows from our earlier work [24], as explained in Sections 8.2–8.3, particularly Lemma 8.3. As a consequence of permanence, the blue trajectory depicted here cannot be a trajectory of (2).

Our explanation of Claim 2.1 involves a "proof by picture". *Concentration space* refers to the space in which the trajectories of the dynamical system (2) evolve, excluding points where X or Y has concentration zero; thus concentration space is $\mathbb{R}^2_{>0}$. *Energy space* has coordinates $u = \log x$ and $v = \log y$, so energy space is (another) $\mathbb{R}^2$. Concentration space and energy space are diffeomorphic via the Lie group isomorphism $(x, y) \mapsto (\log x, \log y)$. This map sends the identity $(1, 1)$ to the origin $(0, 0)$, and the curve parametrized by $\theta \mapsto (\theta^a, \theta^b)$ in concentration space to the ray from the origin in direction (a, b) in energy space:



toric ray in concentration space

a vector in energy space

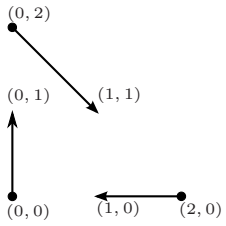
Such curves in concentration space, obtained by exponentiating rays in energy space, are central to our analysis. The curve parametrized by (θ^a, θ^b) is the *toric ray* in direction (a, b).

Remark 2.2. In Lie theory or differential geometry, the inverse map from energy space to concentration space is the usual exponential map from the tangent space at the identity $(1, 1)$ to the group. Thus toric rays are the positive parts of 1-parameter subgroups of concentration space, and they can also be thought of as geodesics.

For every point $(x, y) \in \mathbb{R}^2_{>0}$, the gradient of g at (x, y) is parallel to the direction of the toric ray that passes through that point. This is because $\nabla g(x, y) = (\log x, \log y)$, so that

$$\nabla g(\theta^a, \theta^b) = (\log \theta)(a, b). \tag{3}$$

Combinatorial information about a reaction network is represented in the *space of complexes* by the reaction diagram, which is a geometric representation of the reaction network. The reaction diagram for network (1) is depicted here:



The reaction $pX + qY \xrightarrow{k} rX + sY$, for example, is represented by the arrow from the complex (p, q) to the complex (r, s). (Other authors have sometimes used the term "space of complexes" to refer to a real vector space of dimension equal to the number of vertices of the reaction graph. Readers should be aware that our usage is different. For us, the dimension of the space of complexes equals the number of species.)

The right-hand side of the differential equations (2) consists of a sum in which the summand arising from a reaction of the form $pX + qY \xrightarrow{k} rX + sY$ is:

$$kx^p y^q \begin{pmatrix} r - p \\ s - q \end{pmatrix}. \tag{4}$$

The relation between concentration space and the space of complexes depicted in Figure 1 is central to our analysis. Specifically, along a toric ray parametrized by (θ^a, θ^b), the monomial contribution $x^p y^q$ of the reaction $pX + qY \xrightarrow{k} rX + sY$ equals $\theta^{\langle (a,b),(p,q) \rangle}$. The exponent $\langle (a, b), (p, q) \rangle$ has a geometric interpretation in the space of complexes as the value of the linear functional $ax + by$ at the point (p, q). We now ask: as $\theta \to +\infty$, which reactions are dominant, that is, which reactions yield the largest-magnitude contribution (4)? The answer is that the dominant reactions are the ones with maximal inner product $\langle (a, b), (p, q) \rangle$,

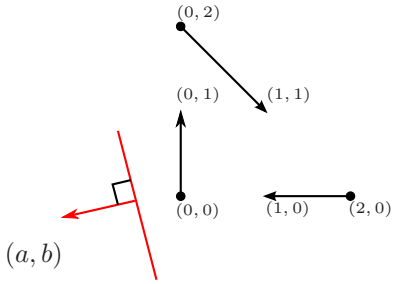


Figure 1: "Proof by picture" for permanence of the reverse Lotka-Volterra system. The toric ray corresponding to the depicted vector pulls hardest on the source of the vertical reaction.

because their monomial contributions grow the fastest as $\theta \to +\infty$ and thus overwhelm the lesser monomial contributions. The directions in which these reactions "pull" (that is, the vector contributions to the differential equations (2)) remain unchanged along the toric ray and are given by their *reaction vectors* $\begin{pmatrix} r-p \\ s-q \end{pmatrix}$ as in (4).

We emphasize this key point: along every toric ray, there exist some asymptotically dominant reactions that determine the dynamics. These reactions can be determined purely combinatorially: they are the reactions $pX+qY \to rX+sY$ whose *reactant complexes* (p,q) attain the maximal inner product with the direction (a,b) of the toric ray.

Returning to Claim 2.1, the time derivative $\frac{d}{dt}g(x(t),y(t))$ of g at a trajectory point at time $t=t_0$ of the form $(x(t_0),y(t_0)) = (\theta^a,\theta^b)$ is determined by the chain rule to be

$$\frac{d}{dt}g(x(t),y(t))\Big|_{t=t_0} = \big\langle \nabla g(\theta^a,\theta^b), P(\theta^a,\theta^b)\big\rangle \tag{5}$$

where $P(x,y)$ denotes the right-hand side of the differential equations (2) evaluated at the point (x,y). By linearity, we can analyze separately the contribution to the derivative (5) of each reaction $pX+qY \xrightarrow{k} rX+sY$. The contribution of such a reaction, from equations (3) and (5), is

$$\langle (a,b),(r-p,s-q)\rangle \theta^{\langle (a,b),(p,q)\rangle} k \log\theta. \tag{6}$$

So far, everything stated above holds for any reaction network with two species, and can be appropriately generalized for more species. Now we appeal to the fact that the reactions in the reaction diagram under consideration "point inward" (i.e., the network is *strongly endotactic*; see Definition 3.14.4).

The term $\log\theta$ in the contribution (6) is common to every reaction, so it is not significant to the analysis of the sign of the derivative $\frac{d}{dt}g(x(t),y(t))$. On the other hand, the reaction

rate constants $k_1, k_2, k_3 \in \mathbb{R}_{>0}$ are significant. However, it turns out that we can ignore the rate constants, and instead simply analyze

$$\text{pull} = \langle (a,b), (r-p, s-q) \rangle \theta^{\langle (a,b), (p,q) \rangle},$$

which we call the *pull* (see Definition 6.18) of the reaction along the toric ray in direction (a,b). Details appear in Sections 7 and 8, particularly in the proof of Theorem 7.5.

Suppose we could verify by inspection of the reaction diagram that for every direction (a,b), the inner products in the pulls of all dominant reactions along toric rays in direction (a,b) satisfy the inequality $\langle (a,b), (r-p, s-q) \rangle < 0$. By compactness of the unit circle, whose points are thought of as directions, we could choose a uniform cutoff θ large enough so that along every toric ray, past this cutoff the monomial contributions of dominant reactions overwhelm the contributions of all other reactions. Indeed, we could ensure this even after including the effects of reaction rate constants. Hence, outside the compact set

$$K_\theta = \left\{ (\theta_0^a, \theta_0^b) \mid (a,b) \text{ is a unit vector in energy space and } \theta_0 \in [1, \theta] \right\} \subseteq \mathbb{R}^2_{>0}$$

in concentration space, the time derivative of g would be negative.

Indeed, in generic directions (a,b), the dominant reactions along the toric ray in direction (a,b) do satisfy the required inequality. However, this fails in precisely three directions, namely, $(-1,0)$, $(0,-1)$, and $(1,1)$.

For instance, consider the direction $(a,b) = (0,-1)$. There exist two dominant reactions: one "sustaining" reaction $0 \xrightarrow{k_2} Y$ with $\langle (a,b), (r-p, s-q) \rangle < 0$, and another reaction $2X \xrightarrow{k_1} X$ with $\langle (a,b), (r-p, s-q) \rangle = 0$. The reaction $2X \xrightarrow{k_1} X$ makes no contribution to the derivative of g in direction $(0,-1)$. However, consider a nearby direction $(-\varepsilon, -1)$ for some small $\varepsilon > 0$. In this direction, the reaction $2X \xrightarrow{k_1} X$ is now "draining": its pull is strictly greater than 0. It is true that if $\varepsilon > 0$ is fixed, then along the toric ray in direction $(-\varepsilon, -1)$, this reaction is eventually dominated by the sustaining reaction, because its monomial term is now smaller. However, as ε gets smaller, the value of the cutoff θ after which this domination of monomials occurs must become arbitrarily large.

This is problematic, because the compact set K_θ requires a single value of θ to work as a cutoff for every direction (a,b). To accomplish this, we turn to a second observation: for sufficiently small $\varepsilon > 0$, the inner product of the reaction vector of $2X \xrightarrow{k_1} X$ with the direction $(-\varepsilon, -1)$ of the toric ray is, although positive, near zero. A more detailed analysis along these lines, using information from both the monomial and the inner product, accomplishes Claim 2.1. This approach is developed in Section 6 via the technology of jets and jet frames. These allow us to reduce the analysis of the sign of the time derivative of g to a combinatorial calculation on the reaction diagram.

Returning to the example, the family of directions $(-\varepsilon, -1)$ as $\varepsilon \to 0^+$ corresponds to the jet frame $\big((0,-1), (-1,0)\big)$. Two reactions dominate along the toric ray in direction $(0,-1)$, namely $2X \to X$ and $0 \to Y$. As for the second direction $(-1,0)$ of the jet frame, only the

sustaining reaction $0 \to Y$ is dominant in that direction. The pull of the draining reaction $2X \to X$ is dominated by the pull of the sustaining reaction $0 \to Y$ in a uniform manner in all directions $(-\varepsilon, -1)$ for small $\varepsilon > 0$, and that produces the required uniform cutoff θ.

In general, for strongly endotactic reaction networks, along every jet frame the pull of each draining reaction is dominated by the pull of some sustaining reaction (Proposition 6.26). This result is key to obtaining a general version of Claim 2.1, which we then apply to prove our main results.

3 Reaction network theory

In this section, we recall the definitions of reaction networks and their associated mass-action differential inclusions, following the notation in our earlier work [24].

3.1 Reaction networks and reaction systems

Definition 3.1. A *reaction network* $(\mathcal{S}, \mathcal{C}, \mathcal{R})$ is a triple of finite sets: a set $\mathcal{S}$ of *species*, a set $\mathcal{C} \subseteq \mathbb{R}^{\mathcal{S}}$ of *complexes*, and a set $\mathcal{R} \subseteq \mathcal{C} \times \mathcal{C}$ of *reactions*. The *reaction graph* is the directed graph $(\mathcal{C}, \mathcal{R})$ whose vertices are the complexes and whose directed edges are the reactions. A reaction $r = (y, y') \in \mathcal{R}$, also written $y \to y'$, has *reactant* $y = \text{reactant}(r) \in \mathbb{R}^{\mathcal{S}}$, *product* $y' = \text{product}(r) \in \mathbb{R}^{\mathcal{S}}$, and *reaction vector*

$$\text{flux}(r) = \text{product}(r) - \text{reactant}(r) = y' - y.$$

A *linkage class* is a connected component of the reaction graph. The *reaction diagram* is the realization $(\mathcal{C}, \mathcal{R}) \to \mathbb{R}^{\mathcal{S}}$ of the reaction graph that takes each reaction $r \in \mathcal{R}$ to the edge from reactant(r) to product(r). The *reactant polytope* is the convex hull $\text{Conv}\{y \in \mathbb{R}^{\mathcal{S}} \mid y \to y' \in \mathcal{R}\}$ of the reactant complexes.

Beginning in the following example, we follow the usual conventions of depicting a network by its reaction graph or reaction diagram and writing a complex as, for example, $2A + B$ rather than $y = (2, 1)$.

Example 3.2. The following network has two species (A and B), five complexes, four reactions (each indicated by a unidirectional arrow), and two linkage classes:

$$2A \rightleftharpoons A + B \qquad B \to 0 \to 2B.$$

The reaction polytope is the convex hull of the four reactant complexes $(2, 0), (1, 1), (0, 1)$, and $(0, 0)$.

Remark 3.3. The chemical reaction network theory literature usually imposes the following requirements for a reaction network.

- Each complex takes part in some reaction: for all $y \in \mathcal{C}$ there exists $y' \in \mathcal{C}$ such that $(y, y') \in \mathcal{R}$ or $(y', y) \in \mathcal{R}$; and
- no reaction is trivial: $(y, y) \notin \mathcal{R}$ for all $y \in \mathcal{C}$.

Definition 3.1 does *not* impose these conditions: our reaction graphs may include isolated vertices or self-loops. In our earlier work, we dropped these conditions to ensure that the projection of a network—obtained by removing certain species—remains a network under our definition even if some reactions become trivial [24]. In addition, like Craciun, Nazarov, and Pantea [15, §7], we allow arbitrary real complexes $y \in \mathbb{R}^{\mathcal{S}}$. Thus our setting is more general than that of usual chemical reaction networks, whose complexes $y \in \mathbb{Z}^{\mathcal{S}}_{\geq 0}$ are nonnegative integer combinations of species, as in the next definition. The ODE systems defined in §3.3 that result from real complexes have been studied over the years and called "power-law systems".

Definition 3.4. A reaction network $(\mathcal{S}, \mathcal{C}, \mathcal{R})$ is

1. *chemical* if $\mathcal{C} \subseteq \mathbb{Z}^{\mathcal{S}}_{\geq 0}$;
2. *reversible* if the reaction graph of the network is undirected: a reaction (y, y') lies in $\mathcal{R}$ if and only if its reverse reaction (y', y) also lies in $\mathcal{R}$;
3. *weakly reversible* if every linkage class of the network is strongly connected.

Definition 3.5. The *stoichiometric subspace* H of a network is the span of its reaction vectors. The *dimension* of a network is the dimension of its stoichiometric subspace H. For a positive vector $x_0 \in \mathbb{R}^{\mathcal{S}}_{>0}$, the *invariant polyhedron* of x_0 is the polyhedron

$$\mathcal{P} = (x_0 + H) \cap \mathbb{R}^{\mathcal{S}}_{\geq 0}.$$

This polyhedron is also referred to as the *stoichiometric compatibility class* in the chemical reaction network theory literature [18].

Example 3.6. Recall the network from Example 3.2 with reactions $2A \rightleftharpoons A + B$ and $B \to 0 \to 2B$. This is a two-dimensional chemical reaction network that is not weakly reversible. For every choice of $x_0 \in \mathbb{R}^2_{>0}$, the corresponding invariant polyhedron is the positive orthant: $\mathcal{P} = \mathbb{R}^2_{\geq 0}$.

Another polyhedron of interest appears in the next definition. For an introduction to polyhedral geometry, we refer the reader to the book by Ziegler [51].

Definition 3.7. For a positive integer $n \in \mathbb{Z}_{>0}$, a *polytope* in $\mathbb{R}^n$ is the convex hull of a finite set of points in $\mathbb{R}^n$. The *reactant polytope* of reaction network $(\mathcal{S}, \mathcal{C}, \mathcal{R})$ is the convex hull of the reactant complexes reactant$(\mathcal{R}) \subseteq \mathbb{R}^{\mathcal{S}}$.

We now turn to the concept of a reaction system.

Definition 3.8. Write $\mathit{CmpctInt} = \big\{[a,b] \mid 0 < a \leq b < \infty\big\}$ for the set of compact subintervals of $\mathbb{R}_{>0}$. Let $(\mathcal{S},\mathcal{C},\mathcal{R})$ be a reaction network. A *tempering* is a map $\kappa : \mathcal{R} \to \mathit{CmpctInt}$ that assigns to each reaction a nonempty compact positive interval. A *confined reaction system* consists of a reaction network $(\mathcal{S},\mathcal{C},\mathcal{R})$, a tempering κ, and an invariant polyhedron $\mathcal{P}$ of the network.

3.2 Strongly endotactic chemical reaction networks

Endotactic chemical reaction networks, a generalization of weakly reversible networks, were introduced by Craciun, Nazarov, and Pantea [15, §4]. We introduced strongly endotactic networks, a subclass of endotactic networks, in [24]. We now recall the definitions.

Definition 3.9. The standard basis of $\mathbb{R}^{\mathcal{S}}$ indexed by $\mathcal{S}$ defines a canonical inner product $\langle\cdot,\cdot\rangle$ on $\mathbb{R}^{\mathcal{S}}$ with respect to which the standard basis is orthonormal. Let $w \in \mathbb{R}^{\mathcal{S}}$.

1. The vector w defines a preorder on $\mathbb{R}^{\mathcal{S}}$, denoted by $\leq_w$, in which

$$y \leq_w y' \;\Leftrightarrow\; \langle w,y\rangle \leq \langle w,y'\rangle.$$

   Write $y <_w y'$ if $\langle w,y\rangle < \langle w,y'\rangle$.

2. For a finite subset $Y \subseteq \mathbb{R}^{\mathcal{S}}$, denote by $\mathrm{init}_w(Y)$ the set of $\leq_w$-maximal elements of Y:

$$\mathrm{init}_w(Y) = \big\{y \in Y \mid \langle w,y\rangle \geq \langle w,y'\rangle \text{ for all } y' \in Y\big\}.$$

3. For a reaction network $(\mathcal{S},\mathcal{C},\mathcal{R})$, the set $\mathcal{R}_w \subseteq \mathcal{R}$ of *w-essential reactions* consists of those whose reaction vectors are not orthogonal to w:

$$\mathcal{R}_w = \big\{r \in \mathcal{R} \mid \langle w, \mathrm{flux}(r)\rangle \neq 0\big\}.$$

4. The *w-support* $\mathrm{supp}_w(\mathcal{S},\mathcal{C},\mathcal{R})$ of the network is the set of vectors that are $\leq_w$-maximal among reactants of w-essential reactions:

$$\mathrm{supp}_w(\mathcal{S},\mathcal{C},\mathcal{R}) = \mathrm{init}_w(\mathrm{reactant}(\mathcal{R}_w)).$$

Remark 3.10. In order to simplify the computations in Section 6, we differ from the usual convention [15, 39], by letting $\mathrm{init}_w(Y)$ denote the $\leq_w$-maximal elements rather than the $\leq_w$-minimal elements. Accordingly, the inequalities in Definition 3.14 are switched, so our definition of endotactic is equivalent to the usual one.

Before presenting Definition 3.14, we provide some underlying geometric intuition, first in terms of 1-dimensional projections (Remarks 3.11 and 3.12) and then via reactant polytopes (Remark 3.13). A third interpretation via jet frames appears later in our work (Lemma 6.22 and Proposition 6.24).

Remark 3.11. For a 1-dimensional network, whose reaction diagram $(\mathcal{C}, \mathcal{R})$ lies on a line in $\mathbb{R}^{\mathcal{S}}$, let $w \in \mathbb{R}^{\mathcal{S}}$ be a nonzero vector that generates the stoichiometric subspace. Every nontrivial reaction $y \to y'$ either *points to the right* (points along w) or *points to the left* (points along $-w$).

1. $(\mathcal{S}, \mathcal{C}, \mathcal{R})$ is endotactic if and only if each nontrivial reaction with a leftmost ($\leq_w$-minimal) reactant points to the right, and each nontrivial reaction with a rightmost ($\leq_w$-maximal) reactant points to the left (see the bottom of Figure 2).

2. A 1-dimensional endotactic network that lies on a line is strongly endotactic if and only if there exists a nontrivial reaction $y \to y'$ (which necessarily points to the right) whose reactant y is a leftmost reactant and additionally there exists a nontrivial reaction $z \to z'$ whose reactant is rightmost. See the bottom of Figure 2.

Remark 3.12. We learned from Craciun and Pantea the following intuition behind Definition 3.14 in terms of 1-dimensional projections. Consider a reaction network $(\mathcal{S}, \mathcal{C}, \mathcal{R})$ and a line generated by a nonzero vector $w \in \mathbb{R}^{\mathcal{S}}$. The orthogonal projection of the reaction diagram $(\mathcal{C}, \mathcal{R})$ onto the line is the reaction diagram of a dimension 0 or 1 network whose complexes are the projections $\langle w, y \rangle w$ for $y \in \mathcal{C}$. The network $(\mathcal{S}, \mathcal{C}, \mathcal{R})$ is endotactic (respectively, strongly endotactic) if and only if for all vectors $w \in \mathbb{R}^{\mathcal{S}}$ that are not orthogonal to the stoichiometric subspace, the projection of the network onto the line generated by w is endotactic (respectively, strongly endotactic). See Figure 2. The dual picture to these ideas was explained in [15, Proposition 4.1] by way of the so-called "parallel sweep test."

Remark 3.13. A second geometric interpretation of the strongly endotactic condition is in terms of the reactant polytope Q, which we recall is the convex hull of reactant$(\mathcal{R})$. We say that a reaction $y \to y'$ *points out of* a set P if the line segment from y to y' intersects P only at the point y. A network is strongly endotactic if and only if (1) no reaction with reactant on the boundary of Q points out of Q, and (2) for all vectors w that are not orthogonal to the stoichiometic subspace, the $\leq_w$-maximal face of Q contains a reactant y such that there exists a nontrivial reaction $y \to y'$ that points out of the face (either along the boundary of Q or into the relative interior of Q). See Examples 3.15 and 3.16.

Definition 3.14. Fix a reaction network $(\mathcal{S}, \mathcal{C}, \mathcal{R})$.

1. The network $(\mathcal{S}, \mathcal{C}, \mathcal{R})$ is *w-endotactic* for some $w \in \mathbb{R}^{\mathcal{S}}$ if

$$\langle w, \text{flux}(r) \rangle < 0$$

   for all w-essential reactions $r \in \mathcal{R}_w$ such that reactant$(r) \in \text{supp}_w(\mathcal{S}, \mathcal{C}, \mathcal{R})$.

2. The network $(\mathcal{S}, \mathcal{C}, \mathcal{R})$ is *W-endotactic* for a subset $W \subseteq \mathbb{R}^{\mathcal{S}}$ if $(\mathcal{S}, \mathcal{C}, \mathcal{R})$ is w-endotactic for all vectors $w \in W$.

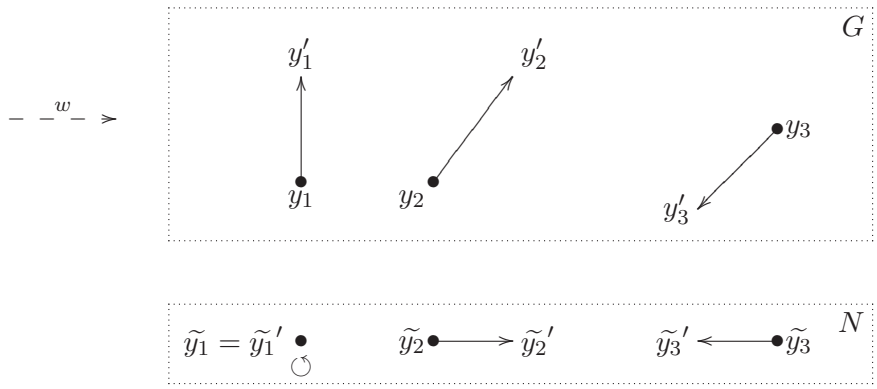


Figure 2: At the top, we depict a direction vector $w = (1, 0)$ and the reaction diagram of a reaction network G with three reactions $y_i \to y_i'$. At the bottom is the reaction diagram of the projection of G to the line generated by w. This 1-dimensional network N is endotactic as explained in Remark 3.11: the leftmost reactant of the nontrivial reactions is $\widetilde{y}_2$ and $\widetilde{y}_2 \to \widetilde{y}_2{}'$ points to the right, and the rightmost reactant complex of the nontrivial reactions is $\widetilde{y}_3$ and $\widetilde{y}_3 \to \widetilde{y}_3{}'$ points to the left. So, G is w-endotactic (but not endotactic: consider the vector $w' = (-1, 1)$). G is not strongly endotactic because $\widetilde{y}_1$ is the unique leftmost reactant of N but $\widetilde{y}_1 \to \widetilde{y}_1{}'$ does not point to the right.

3. The network $(\mathcal{S}, \mathcal{C}, \mathcal{R})$ is *endotactic* if it is $\mathbb{R}^{\mathcal{S}}$-endotactic.

4. $(\mathcal{S}, \mathcal{C}, \mathcal{R})$ is *strongly endotactic* if it is endotactic and for every vector w not orthogonal to the stoichiometric subspace of $(\mathcal{S}, \mathcal{C}, \mathcal{R})$, there exists a reaction $r = (y \to y')$ in $\mathcal{R}$ such that

   (i) $y >_w y'$ (i.e., $\langle w, \text{flux}(r)\rangle < 0$) and

   (ii) y is $\leq_w$-maximal among all reactants in $(\mathcal{S}, \mathcal{C}, \mathcal{R})$: $y \in \text{init}_w(\text{reactant}(\mathcal{R}))$.

Example 3.15. For the network G in Figure 2, the reactant polytope Q is the convex hull of the reactants y_1, y_2, y_3 (labeled by $\bullet$), and both reactions $y_1 \to y_1'$ and $y_3 \to y_3'$ point out of the triangle Q. Thus G is not strongly endotactic.

Example 3.16. The network

$$0 \to 3A + B \qquad 2A \to B \qquad 2B \to A + B$$

is strongly endotactic (thus, endotactic), but not weakly reversible. In light of Remark 3.13, this can be seen from the reaction diagram and reactant polytope Q, which is the convex hull of the reactants $0, 2A, 2B$ (marked by $\bullet$ in Figure 3). Indeed, no reaction points out of the triangle Q, and each proper face of Q—an edge or a vertex—contains at least one reactant in a reaction that points out of that face.

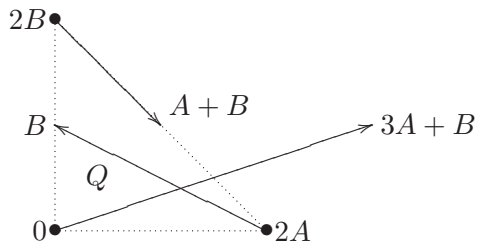


Figure 3: Reaction network from Example 3.16.

Example 3.17. The network from Examples 3.2 and 3.6, whose reactions are $2A \rightleftharpoons A+B$ and $B \to 0 \to 2B$, is endotactic but not strongly endotactic.

The next lemma and the following two corollaries provide examples of strongly endotactic reaction networks. For notation, a set of complexes of a network is a *union of linkage classes* if it is the set of complexes in a union of linkage classes of the reaction graph.

Lemma 3.18. *Let $(\mathcal{S}, \mathcal{C}, \mathcal{R})$ be a weakly reversible reaction network. For a vector $w \in \mathbb{R}^{\mathcal{S}}$, let $T_w = \mathrm{init}_w(\mathrm{reactant}(\mathcal{R}))$ denote the set of $\leq_w$-maximal reactants. Assume that $w \in \mathbb{R}^{\mathcal{S}}$ is orthogonal to the stoichiometric subspace of $(\mathcal{S}, \mathcal{C}, \mathcal{R})$ whenever T_w is a union of linkage classes. Then $(\mathcal{S}, \mathcal{C}, \mathcal{R})$ is strongly endotactic.*

Proof. Assume that $w \in \mathbb{R}^{\mathcal{S}}$ is not orthogonal to the stoichiometric subspace of $(\mathcal{S}, \mathcal{C}, \mathcal{R})$. Every weakly reversible network is endotactic [15, Lemma 4.5], so it remains only to show that there is a reaction going from T_w to the complement $\mathcal{C} \setminus T_w$. All complexes of a weakly reversible network are reactants, so if $y \in T_w$, then $y \geq_w y'$ holds for all complexes y' in $\mathcal{C}$. Thus, by the hypothesis that T_w is not a union of linkage classes, there exists a reaction that goes from T_w to the complement $\mathcal{C} \setminus T_w$ or from $\mathcal{C} \setminus T_w$ to T_w. In the former case, we are done; in the latter case, weak reversibility of the network implies that there is some other reaction that goes from T_w to the complement. Therefore $(\mathcal{S}, \mathcal{C}, \mathcal{R})$ is strongly endotactic. □

We obtain the following corollary for weakly reversible networks.

Corollary 3.19. *If each linkage class of a weakly reversible network $(\mathcal{S}, \mathcal{C}, \mathcal{R})$ has the same stoichiometric subspace, namely that of $(\mathcal{S}, \mathcal{C}, \mathcal{R})$ itself, then $(\mathcal{S}, \mathcal{C}, \mathcal{R})$ is strongly endotactic.*

Proof. Suppose each linkage class has the same stoichiometric subspace H. For $w \in \mathbb{R}^{\mathcal{S}}$ not orthogonal to H, let y be a $\leq_w$-maximal complex, i.e., $y \in T_w = \mathrm{init}_w(\mathrm{reactant}(\mathcal{R}))$. By Lemma 3.18, it suffices to show that T_w is not a union of linkage classes. Letting G_j denote the linkage class of y, it follows that H is spanned by the vectors $z - y$, where z is a

complex in G_j. Thus there exists z^* in G_j such that $y >_w z^*$, because otherwise w would be orthogonal to H. So, $z^* \notin T_w$, which implies that T_w is not a union of linkage classes. □

Corollary 3.19 implies that the networks Anderson treated [2, 4] are strongly endotactic.

Corollary 3.20. *Every weakly reversible reaction network with exactly one linkage class is strongly endotactic.*

3.3 Mass-action differential inclusions

We now recall from [24] how a confined reaction system gives rise to a mass-action differential inclusion. In what follows, we assume that all manifolds have finite dimension.

Definition 3.21. Let M be a smooth manifold with tangent bundle $\pi_M : TM \to M$. A *differential inclusion* on M is a subset $X \subseteq TM$.

Definition 3.22. Let X be a differential inclusion on a smooth manifold M.

1. Let $I \subseteq \mathbb{R}_{\geq 0}$ be a nonempty interval (in particular, connected) containing its left endpoint. A differentiable curve $f : I \to M$ is a *trajectory* of X if the tangent vectors to the curve lie in X.
2. An unbounded interval is a *ray*. A trajectory f defined on a ray *eventually* has property P if there exists $T > 0$ such that property P holds for f whenever $t \geq T$. The *ω-limit set* of a trajectory f defined on a ray is the set

$$\omega(f) = \left\{x \in \overline{M} \mid f(t_n) \to x \text{ for some sequence } t_n \in I \text{ with } t_n \to \infty\right\}$$

of accumulation points of f arising from a sequence of times tending to infinity.

The next definition makes use of the notation $x^y = x_1^{y_1} \cdots x_m^{y_m}$, for $x, y \in \mathbb{R}^m$.

Definition 3.23. The *mass-action differential inclusion* of a confined reaction system, given by a reaction network $(\mathcal{S}, \mathcal{C}, \mathcal{R})$ with tempering κ and invariant polyhedron $\mathcal{P}$, is the differential inclusion on $\mathbb{R}^{\mathcal{S}}_{>0}$ in which the fiber over a point $x \in \operatorname{int}(\mathcal{P})$ is

$$\left\{ \sum_{r \in \mathcal{R}} k_r x^{\text{reactant}(r)} \operatorname{flux}(r) \;\middle|\; k_r \in \kappa(r) \text{ for all } r \in \mathcal{R} \right\} \subseteq \mathbb{R}^{\mathcal{S}} = T_x \mathbb{R}^{\mathcal{S}}_{>0},$$

and the fiber over all other points $x \in \mathbb{R}^{\mathcal{S}}_{>0} \setminus \operatorname{int}(\mathcal{P})$ is empty.

Example 3.24. One possible tempering on the network from Example 3.16 is given by $\kappa(0 \to 3A + B) = [1, 2]$, and $\kappa(2A \to B) = \{3\}$, and $\kappa(2B \to A + B) = [4, 5]$. Every trajectory $x(t) = \big(x_A(t), x_B(t)\big)$ of the resulting mass-action differential inclusion satisfies

$$\begin{aligned} \dot{x}_A &= 3 \cdot k_1(t) - 6 \cdot x_A(t)^2 + k_3(t) \cdot x_B(t)^2 \\ \dot{x}_B &= k_1(t) + 3 \cdot x_A(t)^2 - k_3(t) \cdot x_B(t)^2 \ , \end{aligned}$$

where $k_1(t) \in [1, 2]$ and $k_3(t) \in [4, 5]$ for all time t.

4 Conjectures related to persistence and permanence

In this section, we recall several conjectures related to the persistence of reaction networks. First, some definitions from [24, §2] are required; see [24, §2] for remarks on relations between these various concepts and comparisons with similar notions in the literature.

Definition 4.1. Let $\overline{M}$ be a smooth manifold with corners whose interior is $M = \overline{M} \setminus \partial\overline{M}$, and let $V \subseteq \partial\overline{M}$ be a subset of the boundary. A differential inclusion $X \subseteq TM$ is

1. *persistent* if the closure in $\overline{M}$ of every trajectory of X is disjoint from $\partial\overline{M}$.
2. *repelled by* V if for every open set $O_1 \subseteq \overline{M}$ with $\overline{V} \subseteq O_1$, there exists a smaller open set $O_2 \subseteq O_1$ with $\overline{V} \subseteq O_2$ such that for every trajectory $f : I \to M$ of X, if $f(\inf I) \notin O_1$ then $f(I) \cap O_2$ is empty; in other words, if the trajectory begins outside of O_1, then the trajectory never enters O_2.

If $\overline{M}$ is compact, then a differential inclusion $X \subseteq TM$ is *permanent* if it is persistent and there is a compact subset $\Omega \subseteq M$ such that for every ray I, every trajectory of X defined on I is eventually contained in Ω.

Definition 4.2. A confined reaction system N, specified by a reaction network $(\mathcal{S}, \mathcal{C}, \mathcal{R})$ together with a tempering κ and an invariant polyhedron $\mathcal{P}$, is *persistent* (respectively, *permanent*) if the mass-action differential inclusion on $M = \mathbb{R}^{\mathcal{S}}_{>0}$ arising from the reaction system N is persistent (respectively, permanent) when viewed with respect to the compactification $\overline{M} = [0, \infty]^{\mathcal{S}}$. More generally, a network $(\mathcal{S}, \mathcal{C}, \mathcal{R})$ is itself *persistent* or *permanent* if for all choices of temperings κ and invariant polyhedra $\mathcal{P}$, the resulting mass-action differential inclusion has the corresponding property.

We now state three conjectures, in increasing level of strength, and then state a related fourth conjecture. The first conjecture concerns so-called "complex-balanced" systems, which form a well-studied subclass of weakly reversible mass-action ODE systems. Moreover, this class contains all so-called "detailed-balanced" systems and weakly reversible "deficiency zero" systems. Many properties of complex-balanced systems (as well as detailed-balanced systems and deficiency zero systems) were elucidated by Feinberg, Horn, and Jackson in the 1970s, and we provide only an overview here. (A definition of complex-balanced systems can be found in any of the following references: [13, 19, 26, 27].) For such systems, it is known that a unique steady state resides within the interior of each invariant polyhedron $\mathcal{P}$. This steady state is called the Birch point in [13] due to the connection to Birch's Theorem (see Section 5). Moreover, a strict Lyapunov function exists for this Birch point, so local asymptotic stability relative to $\mathcal{P}$ is guaranteed [27]. An open question is whether all trajectories with initial condition interior to $\mathcal{P}$ converge to the unique Birch point of $\mathcal{P}$. The assertion that the answer is "yes" is the content of the following conjecture, which was stated first by Horn in 1974 [25] and was given the name "Global Attractor Conjecture" by Craciun et al. [13].

Conjecture 4.3 (Global Attractor Conjecture)**.** *For every invariant polyhedron $\mathcal{P}$ of every complex-balanced system, the Birch point $\overline{x} \in \mathcal{P}$ is a global attractor of* $\operatorname{int}(\mathcal{P})$*.*

Due to the strict Lyapunov function, the GAC is equivalent to the following: *every complex-balanced system is persistent.* This suggests the following more general conjecture, which was first stated by Feinberg in 1987 [18, Remark 6.1E].

Conjecture 4.4 (Persistence conjecture)**.** *Every weakly reversible mass-action kinetics ODE system is persistent.*

Conjecture 4.4 was generalized recently by Craciun, Nazarov, and Pantea in the following three ways: the weakly reversible hypothesis is weakened to endotactic, fixed reaction rate constants are allowed to vary within bounded intervals (i.e., they are tempered), and the conclusion of persistence is strengthened to permanence [15, §4].

Conjecture 4.5 (Extended permanence conjecture)**.** *Every endotactic reaction network is permanent.*

Remark 4.6. Conjecture 4.5 captures the intuitively appealing idea that if a reaction diagram "points inwards" then the corresponding dynamics in concentration space must also roughly "point inwards." We interpret this conjecture as a suggestion that the geometry of the reaction diagram ought to be viewed literally as a combinatorial representation of the dynamics. From this perspective, Conjecture 4.5 is a first step in a research program to complete the details of this correspondence. In our previous work [24, Question 5.26], we suggested a framework within which additional aspects of this correspondence might be explored. In particular, we utilized the standard dynamical system notion of topological equivalence to ask if the qualitative nature of the dynamics remains invariant under reasonable transformations to the reaction diagram. This idea was most pithily expressed by asking for the richest domain category (of reaction diagrams) from which the mass-action differential inclusion remains a functor [24].

The following fourth conjecture was stated recently by Anderson [2, §1.1]. It would follow from Conjecture 4.4 according to our definition of persistence for reaction networks (which differs from some other definitions, cf. [24, Remark 2.10]).

Conjecture 4.7 (Boundedness conjecture)**.** *Every weakly reversible mass-action kinetics ODE system has bounded trajectories.*

Although all four conjectures remain open, some progress has been made in recent years. Conjecture 4.7 is true for complex-balanced systems (due to the Lyapunov function: see [47, Lemma 3.5] for details) and was resolved recently for mass-action ODE systems with only one linkage class (under some additional mild hypotheses) by Anderson [2]. Conjecture 4.3 has been proved for systems of dimension at most three [3, 5, 13, 15, 39] and also when the network contains only one linkage class [4]. Conjecture 4.5, and thus Conjectures 4.3

and 4.4 as well, has been resolved in dimensions at most 2 by Craciun, Nazarov, and Pantea for the systems for which Conjecture 4.7 holds [15, 39]. These results are due in part to the analysis of steady states on the boundary of invariant polyhedra.

Indeed, it is known that in the case of complex-balanced systems, the ω-limit set is contained in the set of steady states. In this setting, it has been proved that certain boundary steady states are not ω-limit points. These include vertices of an invariant polyhedron, according to Anderson and Craciun et al. [3, 13]; interior points of facets, according to Anderson and Shiu [5]; and interior points of "weakly dynamically non-emptiable" faces, according to Johnston and Siegel [29]. In fact, trajectories are repelled by such points. Also, some networks have no boundary steady states; for example, Shinar and Feinberg have proved that weakly reversible "concordant" networks have this property [44]. Additionally, the three-dimensional case was resolved by Pantea [39]. The remaining cases for Conjecture 4.3 are systems of dimension 4 and higher in which steady states lie on faces of dimension at least 1 and codimension at least 2.

There have been several other recent approaches to persistence-type results. First, Siegel and Johnston proved that for a complex-balanced system, the positive orthant can be subdivided into strata in which trajectories must obey certain linear Lyapunov functions [46]. Second, Angeli, De Leenheer, and Sontag gave persistence criteria that allow for differential inclusions and time-varying rate constants (i.e., with temperings) [7, 9]. For instance they proved that ω-limit points arising from reaction systems must lie in the relative interior of faces of the positive orthant defined by "critical siphons"; therefore, networks without critical siphons are persistent. (A siphon is a subset of the species whose absence is forward-invariant with respect to the dynamics; see [9, 45] for a precise definition of (critical) siphon.) Another approach makes use of the theory of monotone systems; for instance, see the recent works of Angeli, De Leenheer, and Sontag [8], Banaji and Mierczynski [10], and Donnell and Banaji [17]. Related work has also used the theory of Petri nets [6, 9]. Additionally, Gopalkrishnan proved that every network that violates Conjecture 4.4 must exhibit a certain catalytic property [23]. Finally, we refer the reader to work by Gnacadja [21, 22] that considered a stronger version of persistence, called "vacuous persistence", which allows for trajectories with initial condition on the boundary of an invariant polyhedron as well as in the relative interior; he showed that certain enzymatic networks are persistent in this stronger sense.

For mass-action ODE systems in which persistence is difficult to prove directly, it is possible that the system is dynamically equivalent to—that is, gives rise to the same ODE system as—one that is more easily seen to be persistent. To this end, Szederkényi and Hangos gave a method for determining whether a given system is dynamically equivalent to a complex-balanced or a detailed-balanced one [49]. Similarly, Johnston and Siegel gave algorithms that determine whether a given system is dynamically equivalent—or more generally, is linearly conjugate—to one from certain classes (such as weakly reversible systems) [28, 30].

The results in the current work are complementary to those described above: in Section 8 we resolve Conjectures 4.3–4.7 for strongly endotactic networks in the general setting of

mass-action differential inclusions.

Remark 4.8. The class of complex-balanced systems can be extended in a straightforward way to include some power-law systems, by relaxing the requirement of a complex-balanced system that its complexes be nonnegative integer vectors. In this larger setting, our results concerning the GAC remain valid.

We conclude this section by noting that non-endotactic networks can fail to be persistent; for instance, the network $A \to B$.

Example 4.9. A less trivial example of a non-endotactic non-persistent network is

$$S_0 + F \xrightarrow{\kappa_1} S_1 + E \xrightarrow{\kappa_2} S_2 + F \qquad S_2 \xrightarrow{\kappa_3} S_1 \xrightarrow{\kappa_4} S_0.$$

Write each concentration vector as $x = (x_{S_0}, x_{S_1}, x_{S_2}, x_E, x_F)$. In [6, §IV.A], Angeli showed that if the reaction rate constants satisfy $\kappa_2 < \kappa_4$, then for the resulting mass-action ODE system, the boundary steady state $x^* = (1, 0, 0, 1, 0)$ is locally asymptotically stable relative to its invariant polyhedron. Hence this non-endotactic network is not persistent.

5 Extensions of Birch's theorem

This section extends Birch's theorem in order to deduce Corollary 5.18, which later is used together with Birch's theorem to prove one of our key results, Theorem 7.5. The reader may wish to skip this section on the first pass. The following notation is employed in this section.

Notation 5.1. For vectors $\alpha, \beta \in \mathbb{R}^m_{>0}$, a constant $\theta \in \mathbb{R}_{>0}$, a positive vector $\Theta = (\Theta_1, \ldots, \Theta_m) \in \mathbb{R}^m_{>0}$, and $w = (w_1, \ldots, w_m) \in \mathbb{R}^m$, set

$$\begin{aligned}
\alpha * \beta &= (\alpha_1\beta_1, \ldots, \alpha_m\beta_m) \in \mathbb{R}^m_{>0} \\
\alpha/\beta &= (\alpha_1/\beta_1, \ldots, \alpha_m/\beta_m) \in \mathbb{R}^m_{>0} \\
\theta^w &= (\theta^{w_1}, \ldots, \theta^{w_m}) \in \mathbb{R}^m_{>0} \\
\Theta^w &= \Theta_1^{w_1} \cdots \Theta_m^{w_m} \in \mathbb{R}_{>0} \\
\log(\Theta) &= (\log \Theta_1, \ldots, \log \Theta_m).
\end{aligned}$$

Also, S^{m-1} denotes the unit sphere of dimension $m-1$ in $\mathbb{R}^m$.

Remark 5.2. The top two displayed notations express the group operations in the multiplicative group ("positive real algebraic torus") $\mathbb{R}^m_{>0}$, where $\alpha/\beta = \alpha * \beta^{-1}$.

5.1 Toric rays

Definition 5.3. Let $\alpha \in \mathbb{R}^m_{>0}$.

1. Let $w \in S^{m-1}$. The *toric ray* from α in direction w is the curve
$$\mathrm{Toric}_\alpha(w) = \{\alpha * \theta^w \mid 1 \leq \theta < +\infty\} \subseteq \mathbb{R}^m_{>0}.$$
2. Let $N \subseteq S^{m-1}$. The *toric pencil* $\mathrm{Toric}_\alpha(N)$ from α in directions N is the union of toric rays in directions $w \in N$.

Remark 5.4. The toric ray $\mathrm{Toric}_\alpha(w)$ is the coordinatewise exponential of the infinite ray in $\mathbb{R}^m$ with direction w that originates at the point $\log(\alpha)$.

Remark 5.5. For a given $\alpha \in \mathbb{R}^m_{>0}$, the open rays $\{\log \alpha + (\log \theta)w \mid 1 < \theta < +\infty\}$ partition the punctured space $\mathbb{R}^m \setminus \{\log \alpha\}$. Indeed, when α is the multiplicative identity vector $(1, \ldots, 1)$, so that $\log \alpha$ is the origin, the coordinates $w \in S^{m-1}$ and $\log \theta \geq 0$ may be viewed as polar coordinates on $\mathbb{R}^m$.

Definition 5.6. For $\alpha \in \mathbb{R}^m_{>0}$, the function $g_\alpha : \mathbb{R}^m_{\geq 0} \to \mathbb{R}$ is given by
$$(x_1, \ldots, x_m) \longmapsto \sum_{i=1}^m x_i \log \frac{x_i}{\alpha_i} - x_i,$$
where $0 \log 0$ is defined to be 0. Because $\lim_{x_i \to 0^+} x_i \log x_i = 0$, this definition is the unique choice that makes the function g_α continuous on the closed nonnegative orthant $\mathbb{R}^m_{\geq 0}$, including the boundary.

The next lemma shows that the function g_α captures the geometry of toric rays.

Lemma 5.7. *Let $\alpha \in \mathbb{R}^m_{>0}$, and let $w \in \mathbb{R}^m$ be a unit vector. For every point $x \in \mathrm{Toric}_\alpha(w)$, either $x = \alpha$ or $g_\alpha(x)$ is nonzero and has gradient direction w; that is, $\frac{\nabla g_\alpha(x)}{\|\nabla g_\alpha(x)\|} = w$.*

In words, the lemma says that along a toric ray from α, the direction of the gradient of g_α matches the direction of the toric ray.

Proof. Let α and w be as in the statement of the lemma, and let $\alpha \neq x \in \mathrm{Toric}_\alpha(w)$, so $x = \alpha * \theta^w$ for some $\theta \in (1, \infty)$. A straightforward calculation shows that
$$\nabla g_\alpha(x) = \log(x/\alpha) = (\log \theta)w. \tag{7}$$
Since $\theta \in (1, \infty)$, the coefficient satisfies $\log \theta > 0$. Thus, the gradient of g_α at x is in direction w. Since w is a unit vector, the normalized gradient is $\frac{\nabla g_\alpha(x)}{\|\nabla g_\alpha(x)\|} = w$. □

Remark 5.8. The function g_α has a uniqueness property with respect to toric rays: if f is a function on $\mathbb{R}^m_{>0}$ whose gradient along every toric ray from α points in the direction of that toric ray, then the level sets of f and g_α form the same foliation.

5.2 Outer normal cones

Definition 5.9. Fix a point x of the compactification $[0,\infty]^m$ of $\mathbb{R}^m_{>0}$. Denote by

$$\Sigma_0 = \big\{i \in \{1,\dots,m\} \mid x_i = 0\big\} \quad \text{and} \quad \Sigma_\infty = \big\{i \in \{1,\dots,m\} \mid x_i = \infty\big\}$$

the sets of zero and infinite coordinates of x, respectively. $[0,\infty]^m$ has *outer normal cone*

$$C_x[0,\infty]^m = \{u \in \mathbb{R}^m \mid \operatorname{supp}(u^-) \subseteq \Sigma_0 \text{ and } \operatorname{supp}(u^+) \subseteq \Sigma_\infty\}$$

at x, where $u = u^+ - u^-$ writes u as a difference of nonnegative vectors with disjoint support.

The following lemma states that if a sequence $x(n)$ in $\mathbb{R}^m_{>0}$ converges to a boundary point $x^* \in \partial[0,\infty]^m$, then the limiting direction of $\nabla g_\alpha(x(n))$, if it exists, must lie in the outer normal cone of the hypercube $[0,\infty]^m$ at x^*. It also states a converse.

Lemma 5.10. *Fix a point x^* in the compactification $[0,\infty]^m$ of $\mathbb{R}^m_{>0}$. Let $\alpha \in \mathbb{R}^m_{>0}$ be any positive vector. A unit vector $u \in \mathbb{R}^m$ lies in the outer normal cone of $[0,\infty]^m$ at x^* if and only if there exists a sequence $x(n)$ in $\mathbb{R}^m_{>0}$ such that*

$$x(n) \to x^* \quad \textit{and} \quad \frac{\nabla g_\alpha\big(x(n)\big)}{\|\nabla g_\alpha\big(x(n)\big)\|} \to u. \tag{8}$$

Proof. First suppose that the sequence $x(n)$ for $n \in \mathbb{Z}_{>0}$ satisfies (8). Since $\nabla g_\alpha(x(n)) = \log(x(n)/\alpha)$ by (7), the ith coordinate of $\nabla g_\alpha(x(n))$ goes to $-\infty$ whenever $x_i^* = 0$. Similarly, the ith coordinate of $\nabla g_\alpha(x(n))$ goes to ∞ whenever $x_i^* = \infty$. All other coordinates go to a finite limit, namely $\log\big(x_i^*/\alpha_i\big)$. Therefore, the limit u of the sequence $\frac{\nabla g_\alpha(x(n))}{\|\nabla g_\alpha(x(n))\|}$ must lie in the outer normal cone at x^*.

For the converse, let $u \in \mathbb{R}^m$ be a unit vector in the outer normal cone of $[0,\infty]^m$ at x^*. Define the point $\beta \in \mathbb{R}^m_{>0}$ with coordinates

$$\beta_i = \begin{cases} x_i^* & \text{if } x_i^* \in (0,\infty) \\ 1 & \text{if } x_i^* = 0 \text{ or } x_i^* = \infty. \end{cases}$$

Fix a sequence $\theta(i)$ in $\mathbb{R}_{>1}$ with limit $+\infty$. Consider the sequence of points $x(n) = \beta * \theta(i)^u$ along the toric ray from β in direction u. For $i = 1,\dots,m$,

$$x(n)_i = \begin{cases} x_i^* \cdot \theta(i)^{u_i} & \text{if } 0 < x_i^* < \infty \text{ (thus } u_i = 0) \\ 1 \cdot \theta(i)^{u_i} & \text{if } x_i^* = 0 \text{ (thus } u_i < 0) \\ 1 \cdot \theta(i)^{u_i} & \text{if } x_i^* = \infty \text{ (thus } u_i > 0). \end{cases}$$

Hence, in the first case above, $x(n)_i = x_i^*$ for all n; in the second case, $x(n)_i \to 0 = x_i^*$; and in the third case, $x(n)_i \to \infty = x_i^*$. So, to show (8), it remains only to prove that the normalized gradient of g_α at $x(n)$ converges to u. It is straightforward to verify that

$$\nabla g_\alpha\big(x(n)\big) = \big(\log\theta(i)\big)u + c^*,$$

where c^* is the vector with coordinates $c_i = \log(x_i^*/\alpha_i)$ if $x_i^* \in (0,\infty)$ and 0 if $x_i^* = 0$ or $x_i^* = \infty$. Thus, as $\log\theta(i)$ goes to ∞, the direction of $\nabla g_\alpha(x(n))$ converges to u. □

Lemma 5.11. *Let $H \subseteq \mathbb{R}^m$ be a linear subspace. Let $p \in \mathbb{R}^m_{>0}$. Consider the polyhedron $\mathcal{P} = (p+H) \cap \mathbb{R}^m_{\geq 0}$. Let $x^* \in \partial[0,\infty]^m$. If $x^* \in \overline{\mathcal{P}}$, where the closure is taken in $[0,\infty]^m$, then the outer normal cone of $[0,\infty]^m$ at x^* meets $H^\perp$ only at the origin.*

The key geometric insight for the proof of Lemma 5.11 is that if a translate L' of a support hyperplane L intersects the interior of a polytope P, then L' fails to intersect the support face $L \cap P$ of the untranslated hyperplane, since $\langle u, L'\rangle \neq \langle u, L \cap P\rangle$ for any $u \in L^\perp$.

Proof. Take a sequence $x(n)$ in $\operatorname{int}(\mathcal{P})$ with $\lim_{n\to\infty} x(n) = x^*$, where the limit is taken in the compactification $[0,\infty]^m$, and suppose that u is a unit vector in $C_{x^*}[0,\infty]^m$. At least one of u^+ and u^- is nonzero. If $u^+ \neq 0$, then

$$\langle u, x(n)\rangle \;=\; \langle u^+, x(n)\rangle - \langle u^-, x(n)\rangle \;\geq\; \langle u^+, x(n)\rangle \;\to\; \infty$$

as $n \to \infty$, because the support of u^+ is contained in the set of ∞-coordinates Σ_∞ of x^* and $x(n) \to x^*$. But if $v \in H^\perp$ then $\langle v, x\rangle$ takes a constant finite value for all $x \in \mathcal{P}$; therefore $u \in C_{x^*}[0,\infty]^m \Rightarrow u \notin H^\perp$. In the remaining case, when $u^+ = 0$ and $u^- \neq 0$,

$$\langle u, x(n)\rangle \;=\; -\langle u^-, x(n)\rangle \;\to\; 0$$

as $n \to \infty$, because $\operatorname{supp}(u^-) \subseteq \Sigma_0$. This also prevents $u \in H^\perp$: the inner product $\langle v, x(n)\rangle$ maintains a constant value for all n when $v \in H^\perp$, whereas $\langle u, x(n)\rangle$ is strictly negative and increasing toward 0, because $x(n)$ is a strictly positive vector. □

5.3 Birch's theorem

In Section 5.4, we prove two extensions to the following theorem.

Theorem 5.12 (Birch's theorem)**.** *Let $H \subseteq \mathbb{R}^m$ be a linear subspace, and let $\alpha, p \in \mathbb{R}^m_{>0}$. The relative interior of the polyhedron $\mathcal{P} = (p+H) \cap \mathbb{R}^m_{\geq 0}$ intersects the toric pencil $\operatorname{Toric}_\alpha(H^\perp)$ at exactly one point.*

Variants of Birch's theorem appear across the mathematical sciences. In algebraic statistics, toric pencils appear as log-linear statistical models, and in this setting, Theorem 5.12 was first proved by Birch in 1963 [11]. In dynamical systems, in the setting of chemical reaction systems, the theorem was proved by Horn and Jackson [27, Lemma 4B].

In the setting of quasi-thermostatic chemical reaction systems (a class introduced by Horn and Jackson that includes complex-balanced systems and therefore deficiency zero systems as well [18, 27]), the toric pencil in Birch's theorem is equal to the set of positive steady states. In that context, the toric pencil usually is written as $\{c \in \mathbb{R}^m_{>0} \mid \log\frac{c}{\alpha} \in H^\perp\}$, where α is a given steady state.

Theorem 5.12 prompts the following definition.

Definition 5.13. Let $H \subseteq \mathbb{R}^m$ be a subspace, and let $\alpha, p \in \mathbb{R}^m_{>0}$. The *$\alpha$-Birch point* of the polyhedron $\mathcal{P} = (p + H) \cap \mathbb{R}^m_{\geq 0}$ is the unique point in the intersection $\operatorname{int} \mathcal{P} \cap \operatorname{Toric}_\alpha(H^\perp)$.

We refer to the *Birch point* when the choice of α and $\mathcal{P}$ is clear from context. Horn [25] conjectured that every Birch point of a complex-balanced system is a global attractor of the corresponding invariant polyhedron $\operatorname{int} \mathcal{P}$ (Conjecture 4.3).

The following is a conventional form of Birch's theorem in the setting of algebraic statistics, as stated in the book by Pachter and Sturmfels [37, Theorem 1.10].

Theorem 5.14 (Birch's theorem, restated)**.** *Fix a $d \times m$ real matrix A with columns denoted by $a_1, \ldots, a_m$ and positive vectors $\alpha, p \in \mathbb{R}^m_{>0}$. The image of the monomial map*

$$\begin{aligned} f_{A,\alpha} : \mathbb{R}^d_{>0} &\to \mathbb{R}^m_{>0} \\ \Theta &\mapsto \alpha * (\Theta^{a_1}, \ldots, \Theta^{a_m}) \end{aligned}$$

intersects the polyhedron $\mathcal{P} = \{q \in \mathbb{R}^m_{\geq 0} \mid Aq = Ap\}$ in a single point.

Remark 5.15. To see why Theorems 5.12 and 5.14 are equivalent, first observe that the two definitions of the polyhedron $\mathcal{P}$ coincide because $H = \ker A$. It remains to show that

$$f_{A,\alpha}(\mathbb{R}^d_{>0}) = \operatorname{Toric}_\alpha(H^\perp) = \{\alpha * \theta^w \mid \theta \in \mathbb{R}_{>0} \text{ and } w \in H^\perp\}. \tag{9}$$

Given a vector $\alpha * (\Theta^{a_1}, \ldots, \Theta^{a_m})$ in the image of $f_{A,\alpha}$, when $\theta = \Theta_1$ and

$$w = a(1) + \big(\log_{\Theta_1}(\Theta_2)\big)a(2) + \cdots + \big(\log_{\Theta_1}(\Theta_d)\big)a(d),$$

where $a(j)$ denotes the jth row of A, the point $\alpha * (\Theta^{a_1}, \ldots, \Theta^{a_m}) = \alpha * \theta^w$ lies in the set on the right-hand side of (9). Conversely, given $\alpha * \theta^w$ where $\theta > 0$ and w is a linear combination of the rows of A: $w = c_1 a(1) + \cdots + c_d a(d)$, then for $\Theta = \big(\theta^{c_1}, \ldots, \theta^{c_d}\big)$, it follows that $\alpha * \theta^w = f_{A,\alpha}(\Theta)$ lies in the image of the map $f_{A,\alpha}$. Thus (9) holds.

5.4 Extensions

We extend Birch's theorem to the compactification $[0, \infty]^m$ of the nonnegative orthant $\mathbb{R}^m_{\geq 0}$.

Theorem 5.16 (Birch's theorem at infinity)**.** *Fix a linear subspace $H \subseteq \mathbb{R}^m$ and $\alpha, p \in \mathbb{R}^m_{>0}$. The α-Birch point of the polyhedron $\mathcal{P} = (p+H) \cap \mathbb{R}^m_{\geq 0}$ is the unique point in the intersection*

$$\overline{\mathcal{P}} \cap \overline{\operatorname{Toric}_\alpha(H^\perp)}, \tag{10}$$

where the two closures are taken in the compactification $[0, \infty]^m$.

We call this theorem "Birch's theorem at infinity" because we are studying the same intersection problem as in the original Birch's theorem, but now the parameter θ along toric rays is allowed to take the value $+\infty$.

Proof. The intersection (10) has just one point in $\mathbb{R}^m_{>0}$, namely the α-Birch point of $\mathcal{P}$; this follows from the usual Birch's theorem (Theorem 5.12). Therefore, we need only prove the lack of boundary points x^* in the intersection (10). To this end, let $x^* \in \overline{\text{Toric}_\alpha(H^\perp)} \cap \partial\overline{\mathcal{P}}$ and choose a sequence $x(n) = \alpha * \theta(i)^{w(i)} \to x^*$, where $\theta(i) \to \infty$ and $w(i) \in H^\perp \cap S^{m-1}$. Taking a subsequence if necessary, assume that $w(n)$ converges to some $u \in H^\perp \cap S^{m-1}$. By Lemma 5.7, $\frac{\nabla g_\alpha(x(n))}{\|\nabla g_\alpha(x(n))\|} = w(n) \to u$. By Lemma 5.10, the direction u lies in the outer normal cone of $[0,\infty]^m$ at x^*. Therefore the outer normal cone intersects $H^\perp$ nontrivially, so by Lemma 5.11, the point x^* does not lie in $\partial\overline{\mathcal{P}}$. Thus the intersection (10) is empty. □

By Theorem 5.16, when $w \in H^\perp$ the toric ray $\text{Toric}_\alpha(w)$ either intersects $\overline{\mathcal{P}}$ at the Birch point or does not approach $\overline{\mathcal{P}}$, including its boundary. The following theorem considers toric rays where w lies more generally in a neighborhood of $H^\perp$ and states that these toric rays do not approach $\overline{\mathcal{P}}$ outside a compact neighborhood of the Birch point.

Theorem 5.17 (Perturbed Birch's theorem)**.** *Fix a linear subspace $H \subseteq \mathbb{R}^m$ and vectors $\alpha, p \in \mathbb{R}^m_{>0}$. Let q be the α-Birch point in the polyhedron $\mathcal{P} = (p + H) \cap \mathbb{R}^m_{\geq 0}$, namely the unique point in the intersection* $\text{int}(\mathcal{P}) \cap \text{Toric}_\alpha(H^\perp)$*. For every relatively open neighborhood $\mathcal{O}$ of q in the relative interior of $\mathcal{P}$, the set $H^\perp \cap S^{m-1}$ of unit directions along $H^\perp$ has a relatively open neighborhood N in S^{m-1} such that*

$$(\overline{\mathcal{P}} \setminus \mathcal{O}) \cap \overline{\text{Toric}_\alpha(N)} = \emptyset,$$

where the two closures are taken in the compactification $[0,\infty]^m$.

Proof. For positive integers n, let $\{N_n\}$ denote a shrinking sequence of ε_n-neighborhoods of $H^\perp \cap S^{m-1}$ in S^{m-1} with radius $\varepsilon_n > 0$ tending to 0. (For instance, take $\varepsilon_n = 1/n$.) It suffices to show that the set

$$\mathcal{Q}_n = (\overline{\mathcal{P}} \setminus \mathcal{O}) \cap \overline{\text{Toric}_\alpha(N_n)}$$

is empty for some n. The intersection of the sets $\mathcal{Q}_n$ for all n is $(\overline{\mathcal{P}} \setminus \mathcal{O}) \cap \overline{\text{Toric}_\alpha(H)}$, which is empty by Theorem 5.16. Since the sets $\mathcal{Q}_n$ are nested closed subsets of the compact space $[0,\infty]^m$, a standard theorem of topology (see Theorem 26.9 and the ensuing discussion in the book by Munkres [36]) implies that $\mathcal{Q}_n$ is empty for large n. □

Theorem 5.17 implies that if a toric ray $\text{Toric}_\alpha(w)$ intersects $\mathcal{P}$ outside of a neighborhood of the Birch point q, then w lies outside a neighborhood of $H^\perp$, so the $H^\perp$-component of w is not too dominant. We quantify this via a lower bound on the H-component. For notation, every vector $w \in \mathbb{R}^m$ is uniquely expressible as the sum $w_H + w_{H^\perp}$ of a vector $w_H \in H$ and a vector $w_{H^\perp} \in H^\perp$.

Corollary 5.18. *Fix a linear subspace $H \subseteq \mathbb{R}^m$ and $\alpha, p \in \mathbb{R}^m_{>0}$. For every open subset $\mathcal{O}$ of the interior of the polyhedron $\mathcal{P} = (p+H) \cap \mathbb{R}^m_{\geq 0}$ such that $\mathcal{O}$ contains the α-Birch point, there exists $\mu > 0$ such that $\|w_H\| \geq \mu$ for all unit vectors w satisfying $\alpha * \theta^w \in \text{int}(\mathcal{P}) \setminus \mathcal{O}$ for some $\theta > 1$.*

Proof. By the proof of Theorem 5.17, there exists $\varepsilon > 0$ such that the ε-neighborhood N of $H^\perp \cap S^{m-1}$ in S^{m-1} satisfies $(\overline{\mathcal{P}} \setminus \mathcal{O}) \cap \overline{\mathrm{Toric}_\alpha(N)} = \emptyset$. Now define

$$\mu = \inf \big\{ \|w_H\| \,\big|\, w \in S^{m-1} \setminus N \big\}.$$

Then $\mu > 0$ because it is the infimum of a nonnegative continuous function on a compact set that never takes the value 0 because $w_H = 0 \Rightarrow w \in H^\perp \subseteq N$. If $\alpha * \theta^w \in \mathrm{int}(\mathcal{P}) \setminus \mathcal{O}$, then $w \notin N$ by construction of N, so $\|w_H\| \geq \mu$ by definition of μ. □

Remark 5.19. Corollary 5.18 can be extended so that it also considers the case when the closure in $[0,\infty]^m$ of a toric ray intersects $\overline{\mathcal{P}} \setminus \mathcal{O}$. However, we do not need this extension in the following section.

Remark 5.20. Müller and Regensburger gave an extension of Birch's theorem [35, Proposition 3.9] that is different from those presented here (Theorems 5.16 and 5.17). Their result was used to generalize results about complex-balanced systems to the setting of certain generalized mass-action ODE systems.

6 Jets

This section introduces the technology of jets to relate the combinatorial geometry of reaction diagrams to the dynamics. Using this connection, we prove that for strongly endotactic reaction networks, the "draining" reactions—those that pose a threat to persistence—are dominated by "sustaining" reactions (Proposition 6.26).

6.1 Jet frames, unit jets, and toric jets

Recall the meaning of $\theta^w = \theta^{w_1} \cdots \theta^{w_n}$ from Notation 5.1, where $w \in \mathbb{R}^n$ and $\theta \in \mathbb{R}_{>0}$. Recall also the geometric interpretation, from Remarks 5.4 and 5.5, of w and $\log \theta$ together comprising polar coordinates on $\mathbb{R}^n$.

The following notations are standard.

Notation 6.1. The phrase "for large i" is shorthand for "for all i greater than some fixed constant i_0". For two sequences $(f(i))$ and $(g(i))$ taking values in $\mathbb{R}$,

1. $f(i) = O(g(i))$ if there exists $c > 0$ such that $|f(i)| \leq c|g(i)|$ for large i,
2. $f(i) = \Omega(g(i))$ if there exists $c > 0$ such that $|f(i)| \geq c|g(i)|$ for large i, and
3. $f(i) = \Theta(g(i))$ if there exist $c_1, c_2 > 0$ such that $c_1|g(i)| \leq |f(i)| \leq c_2|g(i)|$ for large i.

Following standard practice, we allow the above notations to appear in expressions and inequalities. For instance, $f(i) \geq g(i) + O(h(i))$ means that $f(i) \geq g(i) + k(i)$ for some $k(i) = O(h(i))$.

Our next definition introduces *jets* and *toric jets*. We first fix an orthonormal basis, a *frame*. Intuitively, a jet consists of a sequence of points that, to first order, are pointing in the direction of the first basis vector, with a small perturbation in the second direction, with an even smaller perturbation in the third direction, and so on. A toric jet is obtained by coordinate-wise exponentiation of a jet.

Toric jets serve as test sequences that allow us to deduce persistence. To prove that a certain function is Lyapunov-like, it is not enough to analyze the function along one direction; perturbations around that direction require exploration, as well. It turns out that the perturbations allowed in toric jets are general enough, while at the same time the gradient of the function $x \log x - x$ along a toric jet is easy to analyze; see the proof of Theorem 7.5.

Definition 6.2. Let $n \in \mathbb{Z}_{>0}$.

1. A *frame* is a list $\overline{w} = (w_1, \ldots, w_\ell)$ of mutually orthogonal unit vectors in $\mathbb{R}^n$.

2. A *jet* is a sequence $(w(i))_{i \in \mathbb{Z}_{>0}}$ of vectors in $\mathbb{R}^n$ in the positive span of some frame $\overline{w}$ in $\mathbb{R}^n$ (i.e., $\langle w(i), w_j \rangle > 0$ for all i and j) such that for $j = 1, \ldots, \ell - 1$, the limit $\lim\limits_{i \to \infty} \frac{\langle w(i), w_j \rangle}{\langle w(i), w_{j+1} \rangle}$ exists and equals $+\infty$. A *unit jet* is a jet that consists of unit vectors.

3. A *toric jet* is a sequence $\big(\theta(i)^{w(i)}\big)_{i \in \mathbb{Z}_{>0}}$, where $(w(i))$ is a unit jet in $\mathbb{R}^n$ and $(\theta(i))$ is a sequence in $\mathbb{R}_{>1}$ with $\lim\limits_{i \to \infty} \theta(i) = +\infty$.

The *jet frame* for the jet $(w(i))$ or toric jet $\big(\theta(i)^{w(i)}\big)$ is $\overline{w}$, and these jets are *framed by* $\overline{w}$.

Remark 6.3. Toric jets have a geometric interpretation: the sequence

$$\big(\log \theta(i)\big) w(i) = (\log \theta(i)\big)\big(\beta_1(n) w_1 + \cdots + \beta_\ell(n) w_\ell\big) \tag{11}$$

is to first approximation a sequence going to infinity in direction w_1 in $\mathbb{R}^n$, which is viewed as "energy space"; the second-order correction is in direction w_2, and so on. Exponentiating the sequence (11) yields the toric jet $\big(\theta(i)^{w(i)}\big)$ in "concentration space" $\mathbb{R}^n_{>0}$.

In the context of a reaction network, if the image of a toric jet lies in a given invariant polyhedron $\mathcal{P}$, then the toric jet necessarily approaches the boundary of $\mathcal{P}$ or is unbounded. More precisely, $\big(\theta(i)^{w(i)}\big)$ approaches the boundary of the closure $\overline{\mathcal{P}}$ of $\mathcal{P}$ in the compactification $[0, \infty]^{\mathcal{S}}$. Our approach to proving persistence is to demonstrate that no toric jet is a sequence of points along a trajectory (see the proof of Theorem 7.5). This strategy is a distillation of the argument employed by Anderson in [4].

Remark 6.4. The concepts of jet frame and unit jet were introduced by Miller and Pak in [34, Definition 4.1] for the purpose of describing the interaction of an expanding wavefront on the boundary of a convex polytope infinitesimally after the wavefront hits a new face. Our definitions are related but not identical to theirs.

Example 6.5. Every sequence of points $\theta(i)^w$ along a toric ray with $\theta(i) \to \infty$ is a special case of a toric jet. Here the jet frame is the singleton list containing only the direction w of the toric ray.

We now give two lemmas concerning unit jets.

Lemma 6.6. *If $(w(i))$ is a unit jet with frame $\overline{w} = (w_1, \ldots, w_\ell)$, then $w(i) \to w_1$ as $i \to \infty$.*

Proof. Let $w(i) = \sum_{j=1}^{\ell} \beta_j(i) w_j$, where $\beta_j(i) > 0$. By definition, $\frac{\beta_j(i)}{\beta_{j+1}(i)} \to \infty$ for $j = 1, \ldots, \ell - 1$. As every $w(i)$ and w_j is a unit vector, it follows that $\beta_1(i) \to 1$, and all other $\beta_j(i) \to 0$ (for $j = 2, \ldots, \ell$). □

Lemma 6.7 (Unit jets are abundant)**.** *Let $n \in \mathbb{Z}_{>0}$. Every infinite sequence of unit vectors in $\mathbb{R}^n$ has an infinite subsequence that is a unit jet.*

Proof. Let $(w(i))$ be an infinite sequence of unit vectors in $\mathbb{R}^n$. Throughout this proof, subsequences of $(w(i))$ are denoted again by $(w(i))$ for ease of notation. We proceed by induction on n.

For $n = 1$, if the sequence $(w(i))$ takes the value $+1$ infinitely often, then choose w_1 to be $+1$. Otherwise $w(i)$ takes the value -1 infinitely often, so choose w_1 to be -1. The required unit jet is the constant subsequence $(w(i) = w_1)$ with frame $\overline{w} = (w_1)$.

Now assume that $n \geq 2$. As the sequence $w(i)$ lies in the unit sphere, which is compact, the sequence must have an accumulation point, which we denote by w_1. Restricting to a subsequence if necessary, assume that $\lim_{i \to \infty} w(i) = w_1$.

Let $\beta_1(i) = \langle w(i), w_1 \rangle$. Then $\lim_{i \to \infty} \beta_1(i) = \langle w_1, w_1 \rangle = 1$. Take a subsequence with $\beta_1(i) > 0$ for all i. Consider the sequence $(w'(i) = w(i) - \beta_1(i) w_1)$.

Case 1. For large i, the sequence $(w'(i))$ is the zero vector. Then for large i, $w(i) = \beta_1(i) w_1$, so this subsequence is a unit jet with frame $\overline{w} = (w_1)$.

Case 2. The sequence $(w'(i))$ is nonzero infinitely often. Then restrict to a subsequence that is always nonzero. By induction, the sequence of unit vectors $w'(i)/||w'(i)||$ in $w_1^\perp \cong \mathbb{R}^{n-1}$ has a subsequence with frame $(w_2, \ldots, w_\ell) \subseteq w_1^\perp$ and corresponding positive coefficients $\widetilde{\beta}_j(i)$ for $j = 2, \ldots, \ell$ such that $w'(i)/||w'(i)|| = \sum_{j=2}^{\ell} \widetilde{\beta}_j(i) w_j$ is a unit jet.

For $j = 2, \ldots, \ell$, let $\beta_j(i) = \widetilde{\beta}_j(i) \cdot ||w'(i)||$. We claim that the corresponding subsequence defined by $w(i) = \beta_1(i) w_1 + \sum_{j=2}^{\ell} \beta_j(i) w_j$ is a unit jet framed by $(w_1, \ldots, w_\ell)$. We need only check that the first coefficient dominates the second. To see that it does, note that $\beta_2(i) = \widetilde{\beta}_2(i) \cdot \|w'(i)\|$ converges to 0, because $\widetilde{\beta}_2(i) \to 1$ by definition of jet frame and $||w'(i)|| \to 0$ by definition of w_1 and β_1. Hence $\beta_1(i)/\beta_2(i) \to \infty$ because $\beta_1(i) \to 1$. □

6.2 Geometry of jets

Theorem 6.11 in this subsection conveys an important geometric idea behind jets. Later we use this result to prove a useful characterization of endotactic networks (Lemma 6.22). We begin with some preliminaries.

Definition 6.8. Fix a positive integer $n \in \mathbb{Z}_{>0}$, a finite set $Q \subseteq \mathbb{R}^n$, and a list $\overline{w} = (w_1, \ldots, w_\ell)$ of vectors in $\mathbb{R}^n$. Set $\mathrm{Super}_0(Q, \overline{w}) = Q$, and inductively for $j = 1, \ldots, \ell$ let $\mathrm{Super}_j(Q, \overline{w})$ denote the set of $\leq_{w_j}$-maximal elements of $\mathrm{Super}_{j-1}(Q, \overline{w})$.

Notation 6.9. When the set Q and the list $\overline{w}$ of vectors is clear from context, we write Super_j to denote $\mathrm{Super}_j(Q, \overline{w})$.

Lemma 6.10. *Fix a positive integer $n \in \mathbb{Z}_{>0}$, a frame $\overline{w} = (w_1, \ldots, w_\ell)$ in $\mathbb{R}^n$, a unit jet $(w(i) = \sum_{j=1}^{\ell} \beta_j(i) w_j)$ framed by $\overline{w}$, and a finite set $Q \subseteq \mathbb{R}^n$. Let $\lambda \in \{1, \ldots, \ell\}$.*

1. *If $x \in \mathrm{Super}_\lambda$ and $y \notin \mathrm{Super}_\lambda$, then there exists $k \in \{1, \ldots, \lambda\}$ such that $\langle w_k, x-y\rangle > 0$ and $\langle w_j, x-y\rangle = 0$ for all $j = 1, \ldots, k-1$. Consequently, $\langle w(i), x-y\rangle > 0$ for large i, and $\langle w(i), x-y\rangle = \Omega(\beta_k(i))$.*

2. *If $\lambda > 1$ and $x \in \mathrm{Super}_{\lambda-1}$ and $y \in Q$, then for large i,*

$$\langle w(i), x-y\rangle \geq \sum_{j=\lambda}^{\ell} \beta_j(i) \langle w_j, x-y\rangle. \tag{12}$$

Proof. 1. Suppose $x \in \mathrm{Super}_\lambda$ and $y \notin \mathrm{Super}_\lambda$. Let $k \in \{1, \ldots, \lambda\}$ be the smallest positive integer such that $y \notin \mathrm{Super}_k$. It follows that $\langle w_k, x-y\rangle > 0$ because $y \in \mathrm{Super}_{k-1} \setminus \mathrm{Super}_k$ and $x \in \mathrm{Super}_k$. Further, $\langle w_j, x-y\rangle = 0$ for all $j < k$, since $x, y \in \mathrm{Super}_j$ for $j < k$. Thus

$$\langle w(i), x-y\rangle \;=\; \sum_{j=1}^{\ell} \beta_j(i) \langle w_j, x-y\rangle \;=\; \beta_k(i)\langle w_k, x-y\rangle + O(\beta_{k+1}(i)),$$

where $\beta_{\ell+1}(i)$ is understood to be 0. Thus, $\langle w(i), x-y\rangle > 0$ for large i because $\langle w_k, x-y\rangle > 0$ and, by definition of unit jet, $\beta_k(i) > 0$ and $\beta_{k+1}(i)/\beta_k(i) \to 0$ as $i \to \infty$. Additionally, it follows that $\langle w(i), x-y\rangle = \Omega(\beta_k(i))$.

2. Suppose $x \in \mathrm{Super}_{\lambda-1}$. Set

$$\phi(i) = \sum_{j=1}^{\lambda-1} \beta_j(i) \langle w_j, x-y\rangle,$$

and rewrite inequality (12) as $\phi(i) \geq 0$. If $y \in \mathrm{Super}_{\lambda-1}$, then $\langle w_j, x-y\rangle = 0$ for $j = 1, \ldots, \lambda-1$, so in fact $\phi(i) = 0$. If $y \notin \mathrm{Super}_{\lambda-1}$, then consider the sequence defined by

$$v(i) = \frac{\sum_{j=1}^{\lambda-1} \beta_j(i) w_j}{||\sum_{j=1}^{\lambda-1} \beta_j(i) w_j||},$$

which is a unit jet framed by $(w_1, \ldots, w_{\lambda-1})$ by construction, using the fact that $(w(i))$ is a unit jet. Then

$$\phi(i) = \Big\| \sum_{j=1}^{\lambda-1} \beta_j(i) w_j \Big\| \langle v(i), x - y \rangle. \tag{13}$$

Of the terms in (13), $\sum_{j=1}^{\lambda-1} \beta_j(i) w_j$ approaches w_1 (by definition of unit jet), so its norm $|| \sum_{j=1}^{\lambda-1} \beta_j(i) w_j ||$ approaches 1, and $\langle v(i), x - y \rangle > 0$ for large i by part 1 of this lemma. So $\phi(i) > 0$ for large i, and inequality (12) holds. □

The next result and Corollary 6.12 may be seen as generalizations of Proposition 2.3(iii) in Ziegler's *Lectures on Polytopes* [51]; with more careful accounting, we obtain our required stronger result. An alternative, direct proof of Corollary 6.12 follows the same lines as the proof in Ziegler's book, without appealing to jets.

Theorem 6.11 (Fundamental theorem of jets)**.** *Fix a positive integer $n \in \mathbb{Z}_{>0}$, a frame $\overline{w} = (w_1, \ldots, w_\ell)$ in $\mathbb{R}^n$, a unit jet $\big(v(i) = \sum_{j=1}^{\ell} \beta_j(i) w_j\big)$ framed by $\overline{w}$, and a finite set $Q \subseteq \mathbb{R}^n$. Then for large i, the $\leq_{v(i)}$-maximal subset of Q equals* Super_ℓ.

Proof. The proof is by induction on ℓ. More precisely, for $j = 1, \ldots, \ell$, let $(v_j(i))$ be the jet

$$v_j(i) = \beta_1(i) w_1 + \cdots + \beta_j(i) w_j.$$

We prove that for all $j = 1, \ldots, \ell$, the $\leq_{v_j(i)}$-maximal subset of Q equals Super_j for large i. The base case of the induction—that Super_1 is the $\leq_{v_1(i)}$-maximal subset of Q for large i—is true by definition of Super_1 because $v_1(i) \to w_1$.

Fix $j < \ell$ and assume that the $\leq_{v_j(i)}$-maximal subset of Q equals Super_j for large i. From now on, consider only such sufficiently large i. For $x \in \mathrm{Super}_j$, let $c(i) = \langle v_j(i), x \rangle$ be the value of the inner product of $v_j(i)$ with every element in Super_j; informally, write $c(i) = \langle v_j(i), \mathrm{Super}_j \rangle$. Similarly, let $c' \in \mathbb{R}$ denote the value of the inner product of w_{j+1} with every element in Super_{j+1}; informally, write $c' = \langle w_{j+1}, \mathrm{Super}_{j+1} \rangle$. Next, define $\delta(i)$ to be the reciprocal of left-hand side of the inequality

$$\frac{1}{\delta(i)} = \max_{y \in Q \setminus \mathrm{Super}_j} \frac{\max\{0, \langle w_{j+1}, y \rangle - c'\}}{c(i) - \langle v_j(i), y \rangle} \leq \frac{\max_{y \in Q \setminus \mathrm{Super}_j} \{0, \langle w_{j+1}, y \rangle - c'\}}{\min_{y \in Q \setminus \mathrm{Super}_j} (c(i) - \langle v_j(i), y \rangle)}. \tag{14}$$

If $1/\delta(i) = 0$, then declare $\delta(i) = +\infty$. Note that $\delta(i) > 0$ for large i, because the denominators in (14) are strictly positive by the inductive hypothesis, and the numerators are nonnegative by construction. The numerator on the right-hand side of (14) has no dependence on i, and hence is $O(1)$. Also, using the identity $\langle v_j(i), x \rangle = c(i)$ for $x \in \mathrm{Super}_j$, the denominator on the right-hand side of (14) satisfies

$$\min_{y \in Q \setminus \mathrm{Super}_j} \big(c(i) - \langle v_j(i), y \rangle\big) = \min_{\substack{y \in Q \setminus \mathrm{Super}_j \\ x \in \mathrm{Super}_j}} \langle v_j(i), x - y \rangle = \Omega\big(\beta_j(i)\big),$$

where the second equality follows from Lemma 6.10.1. Thus the inequality (14) yields

$$\delta(i) \geq \frac{\min_{y\in Q\setminus \mathrm{Super}_j}\big(c(i)-\langle v_j(i),y\rangle\big)}{\max_{y\in Q\setminus \mathrm{Super}_j}\{0,\langle w_{j+1},y\rangle - c'\}} = \frac{\Omega\big(\beta_j(i)\big)}{O(1)} = \Omega\big(\beta_j(i)\big).$$

Hence, by the definition of jet, $\delta(i) > \beta_{j+1}(i)$ for large i.

We now complete the proof by proving that for large i, the following inequality holds for all $y \in Q$ and achieves equality exactly on the set Super_{j+1}:

$$\langle v_{j+1}(i), y\rangle = \langle v_j(i), y\rangle + \beta_{j+1}(i)\langle w_{j+1}, y\rangle \leq c(i) + \beta_{j+1}(i)c'. \tag{15}$$

If $y \in \mathrm{Super}_j$, then $\langle v_j(i), y\rangle = c(i)$ by definition and $\langle w_{j+1}, y\rangle \leq c'$ by construction, with equality holding precisely when $y \in \mathrm{Super}_{j+1}$; hence the desired (in)equality in (15) holds. Now assume that $y \in Q \setminus \mathrm{Super}_j$. There are two subcases based on the sign of $\langle w_{j+1}, y\rangle - c'$. First, if $\langle w_{j+1}, y\rangle - c' < 0$, then the inequality in (15) is strict, noting that $\langle v_j(i), y\rangle \leq c(i)$ because $y \in Q$. In the remaining case, when $y \in Q \setminus \mathrm{Super}_j$ and $\langle w_{j+1}, y\rangle - c' \geq 0$, the construction of $\delta(i)$ in (14) yields the first of the inequalities

$$c(i) - \langle v_j(i), y\rangle \geq \delta(i)\big(\langle w_{j+1}, y\rangle - c'\big) > \beta_{j+1}(i)\big(\langle w_{j+1}, y\rangle - c'\big), \tag{16}$$

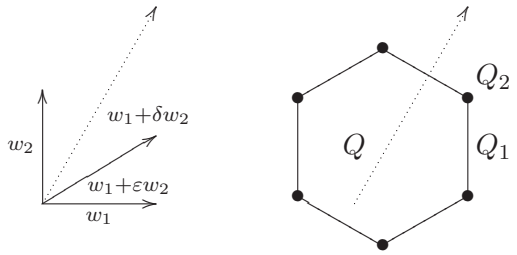
while the second inequality is due to the inequality $\delta(i) > \beta_{j+1}(i)$ proven earlier. Rearranging (16) yields the desired strict inequality in (15). □

Corollary 6.12. *Fix a positive integer $n \in \mathbb{Z}_{>0}$, a frame $\overline{w} = (w_1, \ldots, w_\ell)$ in $\mathbb{R}^n$, and a finite set $Q \subseteq \mathbb{R}^n$. There exist $\beta_1, \ldots, \beta_\ell \in \mathbb{R}_{>0}$ and a positive linear combination $\widetilde{w} = \beta_1 w_1 + \cdots + \beta_\ell w_\ell$ such that the set of $\leq_{\widetilde{w}}$-maximal elements of Q equals Super_ℓ. Additionally, if $\gamma_1, \ldots, \gamma_\ell \in \mathbb{R}_{>0}$, then $\beta_1, \ldots, \beta_\ell$ can be chosen so that $\beta_j < \gamma_j\beta_{j-1}$ for all $j = 2, \ldots, \ell$.*

Proof. Use Theorem 6.11: first consider a jet $(w(i))$ in the frame $\overline{w}$, and then define $\widetilde{w}$ as $w(i_0)$ for a sufficiently large value of i_0, and use the definition of jet. □

Remark 6.13. The hypotheses of Corollary 6.12 can be weakened. The corollary remains true for any set Q whose convex hull is a polytope and for any list of vectors $(w_1, \ldots, w_\ell)$ that need not be orthonormal. Indeed, this is the setting of the analogous result in Ziegler's book [51, Proposition 2.3(iii)]. We do not need this generality here.

Example 6.14. We illustrate Corollary 6.12 by considering the set Q consisting of the six vertices of a hexagon:



The set Q_1 consists of the two vertices on the edge that is the $\leq_{w_1}$-maximal face of the hexagon, and the vertex Q_2 is the $\leq_{w_2}$-maximal subset of Q_1. For $0 < \varepsilon < \delta = \sqrt{3}$, the vertex Q_2 is the $\leq_{w_1+\varepsilon w_2}$-maximal subset of Q. In other words, $w_1 + \varepsilon w_2$ defines Q_2 as long as it lies between w_1 and the dotted arrow depicted twice in the figure above.

6.3 Jets and reactions

One of our primary goals, carried out in Section 7, is to prove that given a strongly endotactic network, there is a compact set outside of which the function $g(x) = \sum_{i \in \mathcal{S}} x_i \log x_i - x_i$ decreases along trajectories; that is, its derivative along trajectories is negative (Theorem 7.5). Intuitively, this derivative is a sum of contributions ("pulls"; Definition 6.18) from each reaction (see the proof of Lemma 7.4). A "draining" reaction acts to hinder our efforts (Definition 6.15) with a positive pull. The pull of a "sustaining" reaction is negative. Proposition 6.26 shows that, after fixing a toric jet, each draining reaction is dominated by some sustaining reaction, so negativity prevails overall and $g(x)$ indeed decreases.

Definition 6.15. For a reaction network $(\mathcal{S}, \mathcal{C}, \mathcal{R})$, let $\overline{w} = (w_1, \ldots, w_\ell)$ be a frame in $\mathbb{R}^{\mathcal{S}}$.

1. A reaction $y \to y' \in \mathcal{R}$ is *$\overline{w}$-essential* if $\langle w_j, y' - y \rangle \neq 0$ for some $j = 1, \ldots, \ell$.
2. The *level* of a $\overline{w}$-essential reaction $y \to y'$ is the least j such that $\langle w_j, y' - y \rangle \neq 0$.
3. A $\overline{w}$-essential reaction $y \to y'$ with level λ is
   (a) *$\overline{w}$-sustaining* if $\langle w_\lambda, y' - y \rangle < 0$.
   (b) *$\overline{w}$-draining* if $\langle w_\lambda, y' - y \rangle > 0$.

Remark 6.16. For a reaction network $(\mathcal{S}, \mathcal{C}, \mathcal{R})$ and a nonzero vector w in $\mathbb{R}^{\mathcal{S}}$, a reaction $y \to y' \in \mathcal{R}$ is (w)-essential (Definition 6.15.1) if and only if $y \to y'$ is w-essential (Definition 3.9.3). More generally, for a frame $\overline{w} = (w_1, \ldots, w_\ell)$ in $\mathbb{R}^{\mathcal{S}}$, a reaction $y \to y'$ is $\overline{w}$-essential if and only if $y \to y'$ is w_j-essential for some $j = 1, \ldots, \ell$.

Example 6.17. For network G of Figure 2 in Section 3, let $\overline{w} = (w_1, w_2)$ be a frame with $w_1 = w$, where w is depicted in Figure 2. Then reaction $y_2 \to y_2'$ is $\overline{w}$-draining, and $y_3 \to y_3'$ is $\overline{w}$-sustaining. Whether $y_1 \to y_1'$ is $\overline{w}$-sustaining or $\overline{w}$-draining depends on whether $\langle w_2, y_1' - y_1 \rangle$ is negative or positive.

Definition 6.18. Let $G = (\mathcal{S}, \mathcal{C}, \mathcal{R})$ be a reaction network.

1. The *pull* of a reaction $y \to y' \in \mathcal{R}$ along a toric jet $\left(\theta(i)^{w(i)}\right)$ in $\mathbb{R}^{\mathcal{S}}$ is the function
$$\text{pull}_{y \to y'}(i) = \langle w(i), y' - y \rangle \theta(i)^{\langle w(i), y \rangle}. \tag{17}$$
2. A unit jet $(w(i))$ in $\mathbb{R}^{\mathcal{S}}$ is *adapted to G* if for every reaction $y \to y' \in \mathcal{R}$, either

(a) $\langle w(i), y' - y\rangle = 0$ for all i; or

(b) $\langle w(i), y' - y\rangle < 0$ for all i, in which case $y \to y'$ is *sustaining along* $(w(i))$; or

(c) $\langle w(i), y' - y\rangle > 0$ for all i, in which case $y \to y'$ is *draining along* $(w(i))$.

A reaction $y \to y'$ is *essential along* $(w(i))$ if it is sustaining or draining along $(w(i))$.

3. A toric jet $\big(\theta(i)^{w(i)}\big)$ framed by $(w_1, \dots, w_\ell)$ in $\mathbb{R}^{\mathcal{S}}$ is *adapted to* G if its unit jet $(w(i))$ is adapted to G and, for all $j = 1, \dots, \ell$, the limit $\lim_{i\to\infty} \theta(i)^{\langle w(i), w_j\rangle}$ exists in $[1, \infty]$.

4. A reaction $r \in \mathcal{R}$ *dominates* a reaction $r' \in \mathcal{R}$ *along* a toric jet $\big(\theta(i)^{w(i)}\big)$ in $\mathbb{R}^{\mathcal{S}}$ if the ratio of their pulls along $\big(\theta(i)^{w(i)}\big)$ tends to infinity in absolute value: $\lim_{i\to\infty} \left|\frac{\mathrm{pull}_r(i)}{\mathrm{pull}_{r'}(i)}\right|$ exists and equals $+\infty$.

When the toric jet is clear from context, we simply say that reaction r *dominates* reaction r'.

Lemma 6.19. *Fix a reaction network $(\mathcal{S}, \mathcal{C}, \mathcal{R})$. Let $w(i)$ be a unit jet in $\mathbb{R}^{\mathcal{S}}$ and let $\big(\theta(i)^{w(i)}\big)$ be a toric jet in $\mathbb{R}^{\mathcal{S}}$.*

1. *There exists an infinite subsequence of $w(i)$ that is adapted to $(\mathcal{S}, \mathcal{C}, \mathcal{R})$.*

2. *There exists an infinite subsequence of $\big(\theta(i)^{w(i)}\big)$ that is adapted to $(\mathcal{S}, \mathcal{C}, \mathcal{R})$.*

Proof. Repeatedly take subsequences, first so that for each reaction $y \to y' \in \mathcal{R}$, the sign of $\langle w(i), y' - y\rangle$ is constant for all i, and then so that for all $j = 1, \dots, \ell$, an accumulation point of $\theta(i)^{\langle w(i), w_j\rangle}$ becomes a limit point. We can do this because $\theta(i) > 1$ and $\langle w(i), w_j\rangle > 0$ for all i, whence $\theta(i)^{\langle w(i), w_j\rangle}$ lies in the compact set $[1, \infty]$. □

It is equivalent for a reaction to be essential along a unit jet or with respect to its jet frame, if the unit jet is adapted to the relevant network. Here is a more precise statement.

Proposition 6.20. *Let $G = (\mathcal{S}, \mathcal{C}, \mathcal{R})$ be a reaction network, and let $(w(i))$ be a unit jet adapted to G and framed by $\overline{w} = (w_1, \dots, w_\ell)$. Let $\lambda \in \{1, \dots, \ell\}$.*

1. *A reaction is $\overline{w}$-sustaining (respectively, $\overline{w}$-draining) if and only if it is sustaining (respectively, draining) along $(w(i))$.*

2. *A $\overline{w}$-essential reaction $y \to y' \in \mathcal{R}$ has level λ if and only if $|\langle w(i), y' - y\rangle| = \Theta\big(\langle w(i), w_\lambda\rangle\big)$.*

Proof. Let $w(i) = \sum_{j=1}^{\ell} \beta_j(i) w_j$. Consider a reaction $y \to y' \in \mathcal{R}$.

If $y \to y'$ is not $\overline{w}$-essential then $\langle w_j, y' - y\rangle = 0$ for all j, and in that case $\langle w(i), y' - y\rangle = 0$ for all i, so $y \to y'$ is not essential along $(w(i))$.

Next assume $y \to y'$ is $\overline{w}$-essential with level λ, so $\langle w_\lambda, y' - y\rangle = 0$ for all $j < \lambda$ whereas $\langle w_\lambda, y' - y\rangle \neq 0$, where the inequality is "$<$" in the sustaining case and "$>$" in the draining

case. Thus $\langle w(i), y' - y\rangle = \beta_\lambda(i)\langle w_\lambda, y' - y\rangle + O(\beta_{\lambda+1}(i))$, so $\langle w(i), y' - y\rangle$ is nonzero for large i by the definition of unit jet—again, with < 0 in the sustaining case and > 0 in the draining case. Hence $y \to y'$ is essential along $(w(i))$ if it is $\overline{w}$-sustaining and draining along $(w(i))$ if it is $\overline{w}$-draining, and in either case $|\langle w(i), y' - y\rangle| = \Theta(\beta_\lambda(i)) = \Theta(\langle w(i), w_\lambda\rangle)$. That takes care of the "$\Rightarrow$" direction of part 2.

For the "$\Leftarrow$" direction of part 2, suppose that the reaction $y \to y'$ is $\overline{w}$-essential and that $|\langle w(i), y' - y\rangle| = \Theta(\beta_\lambda(i))$. It follows that $\langle w_j, y' - y\rangle = 0$ for all $j < \lambda$, and $\langle w_\lambda, y' - y\rangle \neq 0$, so $y \to y'$ has level λ. □

6.4 Jets and endotactic networks

The next two results interpret endotactic and strongly endotactic networks in terms of jets.

Notation 6.21. For a reaction network $(\mathcal{S}, \mathcal{C}, \mathcal{R})$ and jet frame $\overline{w} = \{w_1, \ldots, w_\ell\}$, denote by Super_1 the $\leq_{w_1}$-maximal subset of reactant$(\mathcal{R})$, and for $j \in \{2, \ldots, \ell\}$, write Super_j for the $\leq_{w_j}$-maximal subset of Super_{j-1}.

In other words, the sets Super_j coincide with those in Section 6.2, where $Q = \mathrm{reactant}(\mathcal{R})$.

Lemma 6.22. *Let $(\mathcal{S}, \mathcal{C}, \mathcal{R})$ be a reaction network.*

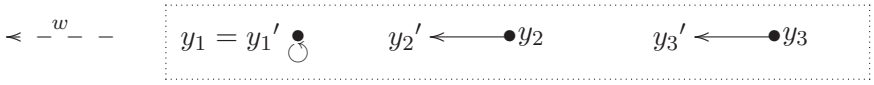
1. *$(\mathcal{S}, \mathcal{C}, \mathcal{R})$ is endotactic if and only if for every singleton frame (w_1) in $\mathbb{R}^{\mathcal{S}}$ and for every (w_1)-draining reaction $y \to y'$ in $\mathcal{R}$, there exists a (w_1)-sustaining reaction $x \to x'$ such that $\langle w_1, x - y\rangle > 0$.*

2. *If $(\mathcal{S}, \mathcal{C}, \mathcal{R})$ is endotactic, then for every frame $\overline{w}$ in $\mathbb{R}^{\mathcal{S}}$ and every $\overline{w}$-draining reaction $y \to y'$ in $\mathcal{R}$ with level λ, the reactant y lies outside of Super_λ.*

Proof. 1. The network $(\mathcal{S}, \mathcal{C}, \mathcal{R})$ fails to be endotactic if and only if there is a unit vector w_1 in $\mathbb{R}^{\mathcal{S}}$ and a reaction $y \to y' \in \mathcal{R}$ with $y \in \mathrm{supp}_{w_1}(\mathcal{S}, \mathcal{C}, \mathcal{R})$ such that $\langle w_1, y' - y\rangle > 0$ (recall Definition 3.14). Equivalently, $(\mathcal{S}, \mathcal{C}, \mathcal{R})$ fails to be endotactic precisely when there is a unit vector w_1 in $\mathbb{R}^{\mathcal{S}}$ and a (w_1)-draining reaction $y \to y' \in \mathcal{R}$ such that every (w_1)-sustaining reaction $x \to x'$ satisfies $y \leq_w x$. Now use that $y \leq_w x \Leftrightarrow \langle w_1, x - y\rangle \leq 0$.

2. Suppose $(\mathcal{S}, \mathcal{C}, \mathcal{R})$ is endotactic. Consider a frame $\overline{w} = (w_1, \ldots, w_\ell)$ in $\mathbb{R}^{\mathcal{S}}$. Suppose $y \to y'$ is $\overline{w}$-draining and has level λ. By Corollary 6.12, Super_λ is the $\leq_{w^*}$-maximal subset of $\mathrm{Super}_0 = \mathrm{reactant}(\mathcal{R})$, where $w^* = w_1 + \varepsilon_1 w_2 + \cdots + \varepsilon_{\lambda-1} w_\lambda$ for some positive numbers $\varepsilon_1, \ldots, \varepsilon_{\lambda-1}$. Corollary 6.12 allows us to choose all of $\varepsilon_1, \ldots, \varepsilon_{\lambda-1}$ to be arbitrarily small, so we can ensure that $\langle w^*, y' - y\rangle$ has the same sign as $\langle w_1, y' - y\rangle$, which is positive. Then, from the definition of w^*-endotactic, we conclude that $y \notin \mathrm{Super}_\lambda$. □

Example 6.23. The converse of Lemma 6.22.2 is false. For the 1-dimensional network



(18)

the essential reactions (if any) corresponding to both (singleton) frames (w) and $(-w)$ are sustaining. More specifically, there are no (w)-essential reactions (although y_1 is the $\leq_w$-maximal reactant, the reaction $y_1 \to y_1'$ is not essential), and the unique $(-w)$-essential reaction $y_3 \to y_3'$ is $(-w)$-sustaining. However, network (18) is not endotactic: $y_2 \to y_2'$ is the leftmost reactant among all of the nontrivial reactions, but it points to the left.

Proposition 6.24. *For an endotactic reaction network $(\mathcal{S}, \mathcal{C}, \mathcal{R})$ with stoichiometric subspace H, the following are equivalent.*

1. *$(\mathcal{S}, \mathcal{C}, \mathcal{R})$ is strongly endotactic.*
2. *For every singleton jet frame (w_1) in $\mathbb{R}^{\mathcal{S}}$ with $w_1 \notin H^\perp$, there exists a (w_1)-sustaining reaction $x \to x'$ such that $x \in \mathrm{Super}_1$.*
3. *For every frame $\overline{w} = (w_1, \ldots, w_\ell)$ in $\mathbb{R}^{\mathcal{S}}$ with $w_1 \notin H^\perp$, there exists a $\overline{w}$-sustaining reaction $x \to x'$ with $x \in \mathrm{Super}_\ell$.*

Proof. The equivalence of items 1 and 2 is straightforward from the definition of strongly endotactic, and item 2 is a special case of item 3. We therefore assume item 1, with the goal of deducing item 3. Let $\overline{w} = (w_1, \ldots, w_\ell)$ be a frame in $\mathbb{R}^{\mathcal{S}}$ with $w_1 \notin H^\perp$. Let $(w(i))$ be a unit jet framed by $\overline{w}$. Use Lemma 6.19.1 to pick a subsequence of $(w(i))$ adapted to $(\mathcal{S}, \mathcal{C}, \mathcal{R})$. Using the fundamental theorem of jets (Theorem 6.11), take i_0 large enough so that the $\leq_{w(i_0)}$-maximal subset of reactant$(\mathcal{R})$ equals Super_ℓ. Taking i_0 even larger, if necessary, assume that $w(i_0) \notin H^\perp$, which is possible because $w(i) \to w_1 \notin H^\perp$. $(\mathcal{S}, \mathcal{C}, \mathcal{R})$ is strongly $w(i_0)$-endotactic, so there exists a $w(i_0)$-sustaining reaction $x \to x'$ with $x \in \mathrm{Super}_\ell$. Since $(w(i))$ is a unit jet adapted to $(\mathcal{S}, \mathcal{C}, \mathcal{R})$, by definition $x \to x'$ is sustaining along the unit jet $(w(i))$. Proposition 6.20.1 implies that $x \to x'$ is $\overline{w}$-sustaining. □

Remark 6.25. Proposition 6.24 can be proven directly, using only Corollary 6.12.

Here is the main result of this section.

Proposition 6.26. *Fix a strongly endotactic reaction network $G = (\mathcal{S}, \mathcal{C}, \mathcal{R})$ with stoichiometric subspace H. Let $\left(\theta(i)^{w(i)}\right)$ be a toric jet adapted to G framed by $\overline{w} = (w_1, \ldots, w_\ell)$ in $\mathbb{R}^{\mathcal{S}}$. If $w_1 \notin H^\perp$ then every draining reaction along $(w(i))$ is dominated by a sustaining reaction along $(w(i))$.*

Proof. Suppose $w_1 \notin H^\perp$. Let $w(i) = \sum_{j=1}^{\ell} \beta_j(i) w_j$. Let $y \to y' \in \mathcal{R}$ be a draining reaction along $(w(i))$. We must find a sustaining reaction $x \to x'$ that dominates $y \to y'$; that is

$$\left| \frac{\mathrm{pull}_{x \to x'}(i)}{\mathrm{pull}_{y \to y'}(i)} \right| = \left| \frac{\langle w(i), x' - x \rangle}{\langle w(i), y' - y \rangle} \right| \theta(i)^{\langle w(i), x - y \rangle} \to \infty. \tag{19}$$

For $k = 1, \dots, \ell$, define a sequence $(v_k(i))$ of unit vectors by

$$v_k(i) = \frac{\sum_{j=1}^{k} \beta_k(i) w_j}{\left\| \sum_{j=1}^{k} \beta_k(i) w_j \right\|}. \tag{20}$$

For $k = 1, \dots, \ell$, the sequence $(v_k(i))$ is a unit jet with frame $(w_1, \dots, w_k)$; this is by construction, starting with the fact that $(w(i))$ is a unit jet.

Let λ be the level of the draining reaction $y \to y'$. By Definition 6.18.3, $L = \lim_{i \to \infty} \theta(i)^{\beta_\lambda(i)}$ exists in $[1, \infty]$. The proof now breaks into two cases based on whether $L = \infty$ or $1 \leq L < \infty$.

Case 1 (Monomial domination): $\lim_{i \to \infty} \theta(i)^{\beta_\lambda(i)} = \infty$.

The network G is strongly endotactic, $w_1 \notin H^\perp$ by hypothesis, and $(w_1, \dots, w_\lambda)$ is a frame in $\mathbb{R}^{\mathcal{S}}$, so by Proposition 6.24 there exists a reaction $x \to x'$ in $\mathcal{R}$ that is $(w_1, \dots, w_\lambda)$-sustaining with $x \in \mathrm{Super}_\lambda$. Hence the level of $x \to x'$ (which exists by definition of sustaining) is at most λ. Also, by Proposition 6.20.1, $x \to x'$ is sustaining along the unit jet $(v_\lambda(i))$. Thus $x \to x'$ is sustaining along $(w(i))$ because $w(i) = \|\sum_{j=1}^{k} \beta_k(i) w_j\| v_k(i) + O(\beta_{k+1}(i))$. Thus it suffices to prove that the limit (19) holds for this reaction $x \to x'$.

Proposition 6.20.2 and the definition of unit jet imply the following asymptotics:

$$\frac{|\langle w(i), x' - x \rangle|}{\beta_\lambda(i)} = \Omega(1) \quad \text{and} \quad \frac{\beta_\lambda(i)}{|\langle w(i), y' - y \rangle|} = \Theta(1).$$

Thus the ratio of inner products satisfies

$$\left| \frac{\langle w(i), x' - x \rangle}{\langle w(i), y' - y \rangle} \right| = \Omega(1). \tag{21}$$

Since $x \in \mathrm{Super}_\lambda$ and $y \in \mathrm{reactant}(\mathcal{R}) \setminus \mathrm{Super}_\lambda$ (by Lemma 6.22.2), they satisfy the hypotheses of Lemma 6.10.1, so we obtain an index $k \leq \lambda$ such that the following inequality holds for the monomial term of interest (which also uses the fact that $\theta(i) > 1$ for all i):

$$\begin{aligned} \theta(i)^{\langle w(i), x - y \rangle} &\geq \theta(i)^{\beta_k(i) \langle w_k, x - y \rangle + O(\beta_{k+1}(i))} \\ &= \big(\theta(i)^{\beta_k(i)}\big)^{\langle w_k, x - y \rangle + O(\beta_{k+1}(i)/\beta_k(i))}. \end{aligned}$$

This last quantity has limit ∞ as i grows because

- $\langle w_k, x - y \rangle > 0$ by definition of the index k that arose from Lemma 6.10.1;
- $\theta(i)^{\beta_k(i)} \geq \theta(i)^{\beta_\lambda(i)} \to \infty$ for large i by assumption in this Case 1; and
- $0 = O\big(\beta_{k+1}(i)/\beta_k(i)\big)$.

Therefore, combining with equation (21), the desired limit (19) holds.

Case 2 (Inner product domination): $1 \le \lim_{i\to\infty} \theta(i)^{\beta_\lambda(i)} < \infty$.

The level of $y \to y'$ satisfies $\lambda > 1$ because otherwise $\lim_{i\to\infty} \theta(i)^{\beta_\lambda(i)} = \lim_{i\to\infty} \theta(i) = \infty$ by Lemma 6.6 and the definition of toric jet. Consider, therefore, the unit jet $(v_{\lambda-1}(i))$ defined in (20) with frame $(w_1, \dots, w_{\lambda-1})$. The network G is strongly endotactic, so as in Case 1, there exists a reaction $x \to x'$ in $\mathcal{R}$, with level $\le \lambda - 1$ and $x \in \mathrm{Super}_{\lambda-1}$, that is sustaining along $(v_{\lambda-1}(i))$ and hence also along $(w(i))$.

Again by Proposition 6.20.2,

$$\lim_{i\to\infty} \frac{|\langle w(i), x' - x\rangle|}{\beta_\lambda(i)} = \infty \quad \text{and} \quad \frac{\beta_\lambda(i)}{|\langle w(i), y' - y\rangle|} = \Theta(1),$$

the former because $\dfrac{|\langle w(i), x' - x\rangle|}{\beta_{\lambda-1}(i)} = \Omega(1)$. Therefore

$$\lim_{i\to\infty} \left| \frac{\langle w(i), x' - x\rangle}{\langle w(i), y' - y\rangle} \right| = \infty. \tag{22}$$

As $x \in \mathrm{Super}_{\lambda-1}$ and $y \in \mathrm{reactant}(\mathcal{R})$, Lemma 6.10.2 implies the first inequality here:

$$\begin{aligned}
\theta(i)^{\langle w(i), x-y\rangle} &\ge \theta(i)^{\beta_\lambda(i)\langle w_\lambda, x-y\rangle + O(\beta_{\lambda+1}(i))} \\
&= \left(\theta(i)^{\beta_\lambda(i)}\right)^{\langle w_\lambda, x-y\rangle + O(\beta_{\lambda+1}(i)/\beta_\lambda(i))} \\
&\to \left(\lim_{i\to\infty} \theta(i)^{\beta_\lambda(i)}\right)^{\langle w_\lambda, x-y\rangle + 0} && \text{because } \beta_\lambda(i) \text{ dominates } \beta_{\lambda+1}(i) \\
&\ge 1 && \text{because } 1 \le \lim_{i\to\infty} \theta(i)^{\beta_\lambda(i)} < \infty.
\end{aligned}$$

Therefore $\theta(i)^{\langle w(i), x-y\rangle} = \Omega(1)$. When combined with (22) this implies that the sustaining reaction $x \to x'$ dominates $y \to y'$. Hence the required limit (19) holds for Case 2. □

Example 6.27. Proposition 6.26 is false without the assumption that $w_1 \notin H^\perp$. For example, consider the network consisting of the single reversible reaction $A \rightleftharpoons B$. The direction $u = (1/\sqrt{2}, 1/\sqrt{2})$ is perpendicular to the stoichiometric subspace, and $v = (-1/\sqrt{2}, 1/\sqrt{2})$ lies in the subspace. Consider an adapted toric jet $\left(\theta(i)^{w(i)}\right)$ framed by (u, v). Then the reaction $B \to A$ is sustaining along $(w(i))$, while the reaction $A \to B$ is draining. The ratio of their pulls is

$$\left| \frac{\mathrm{pull}_{B\to A}(i)}{\mathrm{pull}_{A\to B}(i)} \right| = \theta(i)^{\langle \beta_1(i)u + \beta_2(i)v,\ (0,1)-(1,0)\rangle} = (\theta(i)^{\beta_2(i)})^{\langle v,\ (-1,1)\rangle} = (\theta(i)^{\beta_2(i)})^{\sqrt{2}}. \tag{23}$$

In particular, if $\beta_2(i)$ approaches 0 much faster than $\theta(i)$ approaches ∞, then the limit of $\theta(i)^{\beta_2(i)}$ does not diverge, so the limit of the ratio (23) is not ∞, whence the sustaining reaction does not dominate the draining reaction.

Remark 6.28. The idea of using a sustaining reaction to dominate draining reactions in the proof of Proposition 6.26 is similar to arguments in the proof of Lemma 4.8 in the work of Anderson [4]. Furthermore, Anderson's result is similar to our Theorem 7.5, below: essentially, both results show that a certain function decreases along trajectories outside a compact set. Anderson's concept of "partitioning vectors $y \in \mathcal{C}$ along a sequence of trajectory points" focuses first on the question of which monomials $\theta^{\langle w,y\rangle}$ dominate, and then later in the analysis analyzes the inner product $\langle w, y' - y\rangle$, whereas we consider the entire product $\theta^{\langle w,y\rangle}\langle w, y' - y\rangle$ in the definition of pull (17). Anderson's concept fits into ours in the following way. For a sequence of trajectory points, written as $\theta(i)^{w(i)}$, that do not remain in a compact set, consider the "top tier" defined by the monomial or the pull, respectively. That is, the top tier is the set of reactions $y \to y'$ such that the corresponding monomials or, respectively, corresponding pulls dominate all others along the sequence. For strongly endotactic networks, these two top tiers coincide, with the possible exception of reactions orthogonal to a limiting direction of the vectors $(w(i))$.

Remark 6.29. Some ideas in the proof of Proposition 6.26 are cognate to ideas in Power Geometry [12]. Specifically, in Power Geometry, to determine which terms in a polynomial dominate (for instance, when certain coordinates go to zero of infinity) one works in the log of the coordinates and examines which exponent vectors of the polynomial lie in the relevant face of the Newton polytope of the polynomial.

7 A Lyapunov-like function for strongly endotactic networks

This section uses the results on jets from the previous section to show (Theorem 7.5) that for strongly endotactic networks, outside a compact set the function $g(x) = \sum_{i\in\mathcal{S}} x_i \log x_i - x_i$ from Definition 5.6 decreases along trajectories.

Definition 7.1. If N is a confined reaction system, specified by a reaction network $(\mathcal{S}, \mathcal{C}, \mathcal{R})$, a tempering κ, and an invariant polyhedron $\mathcal{P}$, then g *decreases along trajectories* of the mass-action differential inclusion (Definition 3.23) arising from N *outside a compact set* if there exists a compact set $\mathcal{K} \subseteq \operatorname{int}(\mathcal{P})$ such that for all trajectories $x(t)$, the time derivative satisfies $\frac{d}{dt} g(x(t))|_{t=t^*} < 0$ whenever $x(t^*) \in \operatorname{int}(\mathcal{P}) \setminus \mathcal{K}$.

Remark 7.2. Every strict Lyapunov function decreases outside the compact set consisting of the function's unique minimum. Functions that decrease outside a compact set should be compared with Foster–Lyapunov functions used in the analysis of Markov chains [32, Appendix B.1]. Foster–Lyapunov functions are used, for instance, to prove that a Markov chain always reaches a certain set, the analogue of our compact set $\mathcal{K}$.

The next lemma states that our compact set of interest, $\mathcal{K} = \mathcal{K}_\theta$, is a compact subset of the positive orthant $\mathbb{R}^{\mathcal{S}}_{>0}$. In contrast, some of the sublevel sets of the function g are not compact. Indeed, it is the fact that some level sets of g intersect the boundary of the positive orthant that prevents us from using the sublevel sets as our sets $\mathcal{K}$; see Remark 8.10.

Lemma 7.3. *Let $\mathcal{P}$ be an invariant polyhedron of a reaction network $(\mathcal{S}, \mathcal{C}, \mathcal{R})$, and let $\theta > 1$. Then the set*

$$\mathcal{K}_\theta = \left\{ \theta_0^w \,\middle|\, 1 \leq \theta_0 \leq \theta \text{ and } w \in \mathbb{R}^{\mathcal{S}} \text{ with } ||w|| = 1 \right\} \cap \mathcal{P}$$

is a compact subset of $\operatorname{int}(\mathcal{P})$.

Proof. $\mathcal{K}_\theta$ is compact because it is the intersection of a closed set $\mathcal{P}$ with a continuous image (under coordinatewise exponential) of a compact set, namely, the closed ball of radius $\log \theta$ around the origin in $\mathbb{R}^{\mathcal{S}}$. The intersection is in $\operatorname{int}(\mathcal{P})$ because exponentials never vanish. □

Lemma 7.4. *Let N be a confined reaction system, specified by a network $(\mathcal{S}, \mathcal{C}, \mathcal{R})$, a tempering κ, and an invariant polyhedron $\mathcal{P}$. Suppose there exists $\theta > 1$ such that*

$$\sum_{r=(y \to y') \in \mathcal{R}} k_r \theta_0^{\langle w, y \rangle} \langle w, y' - y \rangle < 0 \tag{24}$$

for all $\theta_0 > \theta$ and unit vectors w in $\mathbb{R}^{\mathcal{S}}$ with $\theta_0^w \in \mathcal{P} \setminus \mathcal{K}_\theta$, and for all $(k_r)_{r \in \mathcal{R}} \in \prod_{r \in \mathcal{R}} \kappa(r)$. Then outside the set $\mathcal{K}_\theta$ in Lemma 7.3, the function $g(x) = \sum_{i \in \mathcal{S}} x_i \log x_i - x_i$ decreases along trajectories of the mass-action differential inclusion (Definition 3.23) arising from N.

Proof. We need $\frac{d}{dt} g(x(t))|_{t=t^*} < 0$ for any trajectory point $x(t^*) \in \operatorname{int}(\mathcal{P}) \setminus \mathcal{K}_\theta$. Such a trajectory point can be written as $x(t^*) = \theta_0^w$ where $\theta_0 > \theta$ and w is a unit vector in $\mathbb{R}^{\mathcal{S}}$, by Lemma 7.3. Additionally, $\frac{d}{dt} x(t)|_{t=t^*}$ has the form given in Definition 3.23 for some rates $k_r \in \kappa(r)$. The following computation is straightforward, using of the gradient of g:

$$\begin{aligned}
\frac{d}{dt} g(x(t))|_{t=t^*} &= \left\langle \nabla g(\theta_0^w), \sum_{r \in \mathcal{R}} k_r (\theta_0^w)^{\text{reactant}(r)} \operatorname{flux}(r) \right\rangle \\
&= \left\langle (\log \theta_0) w, \sum_{r \in \mathcal{R}} k_r \theta_0^{\langle w, \text{reactant}(r) \rangle} \operatorname{flux}(r) \right\rangle \\
&= (\log \theta_0) \sum_{r=(y \to y') \in \mathcal{R}} k_r \theta_0^{\langle w, y \rangle} \langle w, y' - y \rangle.
\end{aligned}$$

As $\log \theta_0 > 0$ because $\theta_0 > 1$, the desired inequality $\frac{d}{dt} g(x(t))|_{t=t^*} < 0$ follows from (24). □

Theorem 7.5. *If N is a confined strongly endotactic reaction system, specified by a reaction network $(\mathcal{S}, \mathcal{C}, \mathcal{R})$, a tempering κ, and an invariant polyhedron $\mathcal{P}$, then outside a compact set $\mathcal{K} \subseteq \mathbb{R}^{\mathcal{S}}_{>0}$, the function $g(x) = \sum_{i \in \mathcal{S}} x_i \log x_i - x_i$ decreases along trajectories of the mass-action differential inclusion (Definition 3.23) arising from N.*

Proof. Begin by fixing a number $0 < \varepsilon < 1$ such that ε is a lower bound for the tempering κ and $1/\varepsilon$ is an upper bound: for all reactions $r \in \mathcal{R}$, $\kappa(r) \subseteq (\varepsilon, 1/\varepsilon)$. The goal is to demonstrate the existence of a cutoff $\theta > 1$ such that g decreases along trajectories outside

of the compact subset $\mathcal{K}_\theta \subseteq \mathrm{int}(\mathcal{P})$ from Lemma 7.3. The proof proceeds by contradiction: assuming that no such θ exists, construct an impossible toric jet.

To be precise, assume $g(x)$ fails to decrease along trajectories outside of each set $\mathcal{K}_\theta$. Pick a sequence $\theta(i) \to \infty$ of real numbers > 1, so $\mathcal{K}_{\theta(1)} \subseteq \mathcal{K}_{\theta(2)} \subseteq \cdots$ and $\bigcup_i \mathcal{K}_{\theta(i)} = \mathbb{R}^{\mathcal{S}}_{>0}$. Lemma 7.4 grants

- a sequence of points $x(i) = \theta(i)^{w(i)} \in \mathcal{P}$, where $w(i)$ is a unit vector in $\mathbb{R}^{\mathcal{S}}$, and
- for each reaction $r \in \mathcal{R}$, a sequence $k_r(i)$ of rates in the interval $\kappa(r)$

such that for all i,

$$\sum_{r=(y\to y')\in\mathcal{R}} k_r(i)\theta(i)^{\langle w(i),y\rangle}\langle w(i), y'-y\rangle \geq 0. \tag{25}$$

Lemma 6.7 produces a subsequence of $(w(i))$ that is a unit jet; call its jet frame $(w_1, \ldots, w_\ell)$ in $\mathbb{R}^{\mathcal{S}}$ and, as usual, denote the subsequence again by $(w(i))$, for ease of notation. The resulting sequence $\big(\theta(i)^{w(i)}\big)$ is a toric jet by construction. Lemma 6.7 affords a subsequence that is adapted to $(\mathcal{S}, \mathcal{C}, \mathcal{R})$, and it is this toric jet $\big(\theta(i)^{w(i)}\big)$ whose impossibility we demonstrate by appeal to Proposition 6.26, which leverages the strongly endotactic hypothesis on $(\mathcal{S}, \mathcal{C}, \mathcal{R})$ to deduce the opposite of the inequality (25).

Both $k_r(i)$ and $\theta(i)^{\langle w(i),y\rangle}$ in (25) remain strictly positive. The only negative contributions to the sum in (25) are from reactions that are sustaining along $(w(i))$, whereas draining reactions contribute positively. To contradict (25), it suffices to show that for any reaction $y \to y'$ that is draining along $(w(i))$, the contribution to the sum (25) of some sustaining reaction $x \to x'$ eventually dominates that of $y \to y'$ by at least a factor of $|\mathcal{R}|$, the number of reactions of the network. In other words, for a fixed draining reaction $r = (y \to y')$, it suffices to exhibit a sustaining reaction $r_{\mathrm{sus}} = (x \to x')$ such that the inequality

$$\frac{k_{r_{\mathrm{sus}}}(i)}{k_r(i)}\theta(i)^{\langle w(i),x-y\rangle}\frac{\langle w(i), x'-x\rangle}{\langle w(i), y-y'\rangle} \;>\; |\mathcal{R}| \tag{26}$$

holds for large i. The number ε was constructed early in the proof so that $k_{r_{\mathrm{sus}}}(i) > \varepsilon$ and $k_r(i) < 1/\varepsilon$. Consequently, the desired inequality (26) follows from the inequality

$$\left|\frac{\mathrm{pull}_{x\to x'}(i)}{\mathrm{pull}_{y\to y'}(i)}\right| \;=\; \theta(i)^{\langle w(i),\; x-y\rangle}\;\frac{\langle w(i), x'-x\rangle}{\langle w(i), y-y'\rangle} \;>\; \frac{|\mathcal{R}|}{\varepsilon^2}.$$

It is sufficient (but not necessary) to prove that the left-hand side of this inequality has limit ∞, i.e., the draining reaction is dominated by some sustaining reaction along this adapted toric jet. This follows from Proposition 6.26, completing the proof, once w_1 is verified not to be orthogonal to the stoichiometric subspace H.

To prove that $w_1 \notin H^\perp$, first express each vector w_j of the jet frame uniquely as a sum $w_{j,H} + w_{j,H^\perp}$ of a vector $w_{j,H} \in H$ and a vector $w_{j,H^\perp} \in H^\perp$. Birch's theorem (Theorem 5.12) produces a unique point q in the intersection $\mathrm{int}(\mathcal{P}) \cap \{\theta^w \mid \theta \in \mathbb{R}_{>0} \text{ and } w \in H^\perp\}$. Pick a neighborhood $\mathcal{O}$ in $\mathrm{int}(\mathcal{P})$ around q whose closure $\overline{\mathcal{O}}$ is also contained in $\mathrm{int}(\mathcal{P})$. The

sequence $\left(\theta(i)^{w(i)}\right)$ eventually avoids $\mathcal{O}$, because $x(i)$ approaches the boundary of the closure $\overline{\mathcal{P}}$ of $\mathcal{P}$ in the compactification $[0,\infty]^{\mathcal{S}}$ (see Remark 6.3). Thus, by Corollary 5.18, there exists a number $0 < \mu \leq 1$ such that the H-components $w_H(i)$ of the unit vectors $w(i)$ have norm at least μ, for large i:

$$\|w_H(i)\| = \|\beta_1(i)w_{1,H} + \cdots + \beta_\ell(i)w_{\ell,H}\| \geq \mu.$$

Since the first coefficient $\beta_1(i)$ approaches 1 while all others approach 0, the left-hand side of this inequality has limit $\|w_{1,H}\|$, so $\|w_{1,H}\| \geq \mu > 0$. Thus $w_{1,H} \neq 0$, as desired. □

Remark 7.6. We wish to emphasize that our extensions of Birch's Theorem (in particular, Corollary 5.18) were used in the proof of Theorem 7.5 to show that $w_1 \notin H^\perp$. This novel use of Birch's Theorem is one of our main contributions.

8 Main results: persistence and permanence

In this section, we prove that strongly endotactic networks have bounded trajectories (Theorem 8.1), are persistent (Theorem 8.5), and are permanent (Theorem 1.1). Although the permanence result is stronger than the first two, the two weaker results are applied in the proof of permanence. In this section, we rely on two key prior results:

- the decrease of the pseudo-Helmholtz free energy function along trajectories outside a compact set (Theorem 7.5), and
- a projection argument from our earlier work [24].

Additionally, we prove the existence of steady states for strongly endotactic networks in the mass-action ODE setting.

Throughout this section, Notation 5.1 remains in effect, as does Definition 5.6, which includes the special case $g(x) = g_{(1,\ldots,1)}(x) = \sum_{i=1}^{|\mathcal{S}|} x_i \log x_i - x_i$ from [19, 26, 27].

8.1 Boundedness of trajectories

Conjecture 4.7 holds for strongly endotactic networks.

Theorem 8.1. *Strongly endotactic networks have bounded trajectories. More precisely, the image of every trajectory of every mass-action differential inclusion (Definition 3.23) arising from a confined strongly endotactic reaction system is bounded.*

Proof. For a mass-action differential inclusion of a confined strongly endotactic reaction system, specified by $(\mathcal{S}, \mathcal{C}, \mathcal{R})$, a tempering κ, and an invariant polyhedron $\mathcal{P}$, let $x : I \to \operatorname{int}(\mathcal{P})$ denote a trajectory. By Theorem 7.5, there exists a compact set $\mathcal{K} \subseteq \operatorname{int}(\mathcal{P})$ such that $\frac{d}{dt} g(x(t))|_{t=t^*} < 0$ for all trajectory points $x(t^*) \in \operatorname{int}(\mathcal{P}) \setminus \mathcal{K}$. Let $M_1 = \sup\{g(x) \mid x \in \mathcal{K}\}$ be the supremum of g on $\mathcal{K}$, which is finite by compactness. Let M be the maximum of M_1

and $g(\inf I)$. Then $x(t)$ remains in the bounded sublevel set $\mathcal{Q} = \{x \in \mathcal{P} \mid g(x) \leq M\}$. Indeed, the trajectory begins in $\mathcal{Q}$ by construction; when it is in $\mathcal{Q}\backslash\mathcal{K}$, the value of g decreases; and while it is in $\mathcal{K}$, the value of g can not exceed M_1 and thus can not exceed M. □

8.2 Persistence

Our next main result is the persistence of strongly endotactic networks (Theorem 8.5). The proof appeals to the following lemma, a special case of [3, Theorem 3.7] of Anderson or [13, Proposition 20] of Craciun et al. It again concerns the function $g(x) = \sum_{i=1}^{m} x_i \log x_i - x_i$, which extends continuously to the boundary of the nonnegative orthant.

Lemma 8.2. *The function $g(x) = \sum_{i=1}^{m} x_i \log x_i - x_i$ defined on $\mathbb{R}^m_{\geq 0}$ has a local maximum at the origin.*

Another result crucial to our proof of Theorem 8.5 is the next lemma, which follows from our earlier work on "vertexical" families of differential inclusions on hypercubes that are well-behaved under maps [24].

Lemma 8.3. *Let $\mathcal{F}$ be the class of all confined strongly endotactic reaction systems. If*

1. *every mass-action differential inclusion of every reaction system in $\mathcal{F}$ is repelled by the origin (Definition 4.1), and*
2. *every trajectory of such a differential inclusion is bounded,*

then every strongly endotactic reaction network is persistent.

Proof. This is a special case of Corollary 6.4 of [24]. □

To be useful, Lemma 8.3 requires the following auxiliary result.

Lemma 8.4. *If the function $g(x) = \sum_{i \in \mathcal{S}} x_i \log x_i - x_i$ for a confined reaction system N decreases along trajectories outside a compact set, then the mass-action differential inclusion (Definition 3.23) arising from N is repelled by the origin.*

Proof. Let O_1 be a relatively open neighborhood of the origin $\overline{0} \in \mathbb{R}^{\mathcal{S}}_{\geq 0}$, and let $\mathcal{P}$ be an invariant polyhedron of $(\mathcal{S}, \mathcal{C}, \mathcal{R})$. Let $\mathcal{K} \subseteq \text{int}(\mathcal{P})$ denote a compact set outside of which g decreases along trajectories. If $\mathcal{P}$ does not contain $\overline{0}$, then all trajectories avoid a fixed neighborhood of the origin because they remain in $\mathcal{P}$. Therefore assume $\mathcal{P}$ contains $\overline{0}$. Let N_ε denote the intersection of the nonnegative orthant $\mathbb{R}^{\mathcal{S}}_{\geq 0}$ with an open ball of radius ε around the origin. Let $\varepsilon > 0$ be sufficiently small so that

- O_1 contains N_ε,
- ε is less than the distance between the origin $\overline{0}$ and the compact set $\mathcal{K}$, and
- the maximum of g on N_ε is attained uniquely at the origin (cf. Lemma 8.2).

As $g(\overline{0}) = 0$, it follows that g is strictly negative on N_ε. Define $-M$ to be the maximum value of g for nonnegative vectors of norm ε, which is attained by compactness:

$$-M = \max\left\{g(z) \,\middle|\, z \in \mathbb{R}^{\mathcal{S}}_{\geq 0} \text{ with } \|z\| = \varepsilon\right\}. \tag{27}$$

Thus, $-M < g(\overline{0}) = 0$. Again by Lemma 8.2, there exists $\delta > 0$ so that $g(z) > -M/2$ for $z \in N_\delta$. It is enough to show that any trajectory $x : I \to \text{int}(\mathcal{P})$ that begins outside O_1, and thus outside N_ε as well, never enters N_δ. Indeed, if the trajectory ever enters the closure of N_ε at some point $x(t^*)$, then $g(x(t^*)) \leq -M$ by (27). Thus, since $g(x(t))$ decreases while $x(t)$ is in N_ε, because $N_\varepsilon \cap \mathcal{K}$ is empty, $x(t)$ never reaches N_δ by definition of δ. □

We can now prove that Conjecture 4.4 holds for strongly endotactic networks networks.

Theorem 8.5. *Strongly endotactic networks are persistent.*

Proof. Follows from Lemmas 8.3 and 8.4 and Theorems 7.5 and 8.1. □

In addition, the GAC holds for strongly endotactic complex-balanced systems.

Proof of Theorem 1.2. Use Theorem 8.5 and the discussion before Conjecture 4.4. □

Remark 8.6. Our reliance on the projection argument (Lemma 8.3) requires showing that for strongly endotactic networks, the origin is repelling and trajectories are bounded. Instead of Lemma 8.3, we could have appealed to a related result, Corollary 6.3 in [24], which states that for a "vertexical" family of differential inclusions—such as the family arising from strongly endotactic networks—if all vertices of $[0,\infty]^{\mathcal{S}}$ are repelling, then the boundary of $[0,\infty]^{\mathcal{S}}$ is repelling and hence these networks are persistent. However, it is not known whether non-origin vertices of $[0,\infty]^{\mathcal{S}}$ are repelling for strongly endotactic networks.

8.3 Permanence

Lemma 8.3 only records half of Corollary 6.4 of [24]. The other half deals with permanence.

Lemma 8.7. *In Lemma 8.3, assume that 1 and 2 hold, and that X is such a differential inclusion on $\mathbb{R}^{\mathcal{S}}_{>0}$. If $K \subseteq \mathbb{R}^{\mathcal{S}}_{>0}$ is a compact set for which there exists $A \in \mathbb{R}_{>0}$ such that every trajectory of X that starts in K remains bounded above by A in each coordinate for all time, then for some compact set $K^+ \subseteq \mathbb{R}^{\mathcal{S}}_{>0}$, no trajectory of X that begins in K leaves K^+.*

Applying Lemma 8.7 requires an auxiliary result that makes use of the persistence result.

Lemma 8.8. *Let X be the mass-action differential inclusion of a confined strongly endotactic reaction system N, specified by a reaction network $(\mathcal{S}, \mathcal{C}, \mathcal{R})$ together with a tempering κ and an invariant polyhedron $\mathcal{P}$. Then there exists a compact set K in* $\text{int}(\mathcal{P})$ *such that the image of every trajectory of X defined on a ray meets K.*

Proof. By Theorem 7.5, there exists a compact set $\mathcal{K} \subseteq \operatorname{int}(\mathcal{P})$ outside of which g decreases along trajectories. Enlarge $\mathcal{K}$ to a closed ε-neighborhood of $\mathcal{K}$, denoted by $\widetilde{\mathcal{K}}$, so that $\widetilde{\mathcal{K}} \subseteq \operatorname{int}(\mathcal{P})$; thus $\widetilde{\mathcal{K}}$ is compact. Let $x : I \to \operatorname{int}(\mathcal{P})$ be a trajectory of the mass-action differential inclusion arising from $(\mathcal{S}, \mathcal{C}, \mathcal{R})$ with κ and $\mathcal{P}$, where $I = [a, \infty]$ is a ray. The goal is to show that the image of x meets $\widetilde{\mathcal{K}}$. If $x(a) \in \widetilde{\mathcal{K}}$, then we are done. Otherwise, there exists a nonempty maximal subinterval $[a, b) \subset I$ on which $g(x(t))$ is decreasing. If $b < \infty$, then $x(b) \in \mathcal{K}$, so we are done. The remaining case has $g(x(t))$ decreasing for all time $t > a$. The closure of the trajectory $x(t)$ is a compact subset of $\operatorname{int}(\mathcal{P})$, because this network has bounded trajectories (Theorem 8.1) and is persistent (Theorem 8.5).

Any trajectory $x(t)$ that fails to intersect $\widetilde{\mathcal{K}}$ is bounded away from $\mathcal{K}$. The time derivative $\frac{d}{dt} g(x(t))$ of such a trajectory is therefore negative and bounded away from 0 (otherwise the closure of the trajectory would meet $\mathcal{K}$, forcing the trajectory itself to meet $\widetilde{\mathcal{K}}$), causing $g(x(t))$ to have limit $-\infty$ as $t \to \infty$. But g is bounded below by $-|\mathcal{S}|$, so we conclude that every trajectory $x(t)$ must visit the compact set $\widetilde{\mathcal{K}}$. □

We now prove our main result, which states that strongly endotactic networks are permanent, thereby resolving Conjecture 4.5 in the strongly endotactic case.

Proof of Theorem 1.1. First, strongly endotactic networks are persistent by Theorem 8.5. Next, by Lemma 8.8, for every mass-action differential inclusion X arising from a strongly endotactic network, there exists a compact set K in the positive orthant that meets every trajectory of X defined on a ray. For the purpose of studying eventual properties of trajectories such as permanence, we therefore assume that each trajectory defined on a ray begins in K. Lemma 8.7 applies for strongly endotactic networks by Lemma 8.4 along with Theorems 7.5 and 8.1; it implies that there exists a compact set K^+ in $\mathbb{R}^{\mathcal{S}}_{>0}$ such that every such trajectory eventually remains in K^+. Thus X is permanent. □

As a corollary, certain weakly reversible networks are permanent.

Corollary 8.9. *If each linkage class of a weakly reversible network $(\mathcal{S}, \mathcal{C}, \mathcal{R})$ has the same stoichiometric subspace, namely that of $(\mathcal{S}, \mathcal{C}, \mathcal{R})$, then $(\mathcal{S}, \mathcal{C}, \mathcal{R})$ is permanent.*

Proof. Follows from Corollary 3.19 and Theorem 1.1. □

Similarly, weakly reversible reaction networks with one linkage class are permanent.

Proof of Theorem 1.3. Follows from the Corollary 3.20 and Theorem 1.1. □

Remark 8.10. Is an appeal to our earlier work (Lemmas 8.3 and 8.7) necessary to prove permanence? Could the fact that g decreases along trajectories suffice to deduce Theorem 1.1? Such an approach seems not to work. On one hand, we can prove that maximal trajectories must approach a prescribed compact set in $\mathcal{P}$, namely the smallest sublevel set Q of g that contains $\mathcal{K}$. On the other hand, Q may meet the boundary of $\mathcal{P}$. Thus the

union of all ω-limit sets of all possible trajectories lies in $Q \cap \operatorname{int}(\mathcal{P})$, but its closure in $\mathcal{P}$ could meet the boundary of $\operatorname{int}(\mathcal{P})$. Another approach to permanence is to determine whether $\mathcal{K}$ is an attracting set. However, this also seems not to work: although persistence can be used to prove that every trajectory outside $\mathcal{K}$ must enter a neighborhood of $\mathcal{K}$ (denoted by $\widetilde{\mathcal{K}}$ in the proof of Lemma 8.8), the trajectory could exit and re-enter $\widetilde{\mathcal{K}}$ infinitely often.

8.4 Existence of positive steady states

In the setting of mass-action ODE systems, Deng et al. proved that for weakly reversible networks, each invariant polyhedron contains a positive steady state [16]. As for two-dimensional endotactic networks, Craciun, Nazarov, and Pantea [15, §6] explained that their permanence result, together with the Brouwer fixed-point theorem [36, §55], implies the existence of positive steady states. (Their result is complementary to that of Deng et al., because weakly reversible networks form a proper subset of endotactic networks.) This standard application of the Brouwer fixed-point theorem to establish the existence of positive steady states in the setting of reaction systems goes back at least to Wei in 1962 [50].

We too are able to prove the existence of positive steady states (Corollary 8.12). However, this is not accomplished via the Brouwer fixed-point theorem, because we do not readily obtain a forward-invariant set U in $\mathbb{R}^{\mathcal{S}}_{>0}$ that is homeomorphic to a ball; in particular, the sets K and K^+ in the proofs of Lemma 8.8 and Theorem 1.1 do not suffice. Instead our proof relies on the following result concerning fixed points of dynamical systems, a special case of [42, Theorem 3] first proven in a general setting by Srzednicki [48].

Lemma 8.11. *Let X be a continuous ODE system on a polyhedron $\mathcal{P}$ in $\mathbb{R}^n$ such that every trajectory of X defined on a ray meets a compact set K. Then X has a steady state.*

Corollary 8.12. *For a mass-action kinetics ODE system arising from a strongly endotactic network, each invariant polyhedron contains a positive steady state.*

Proof. Let X denote such a mass-action kinetics ODE system with invariant polyhedron $\mathcal{P}$. The proof of Lemma 8.8 shows that there exists a compact set K in $\operatorname{int}(\mathcal{P})$ such that the image of every trajectory of X defined on a ray meets K. Thus, by Lemma 8.11, X has a steady state in K, which is therefore in $\operatorname{int}(\mathcal{P})$. □

9 Examples

The examples here illustrate our results as well as their limitations.

9.1 Strongly endotactic networks

Example 9.1. By Theorem 1.1, the strongly endotactic network from Example 3.16, comprised of the reactions $0 \to 3A + B$, $2A \to B$, and $2B \to A + B$, is permanent.

Permanence of the network in Example 9.1 already was known by the results of Craciun, Nazarov, and Pantea [15] and extensions by Pantea [39], which apply to endotactic networks with stoichiometric subspace of dimension at most 2. For the following 3-dimensional example, their results do not apply.

Example 9.2. The following network is strongly endotactic, so it is permanent:

$$0 \to A \to B \to C \to 0.$$

The weaker persistence result for the network in Example 9.2 follows from work of Angeli, De Leenheer, and Sontag [9] because it has no siphons. In contrast, the following example contains critical siphons, so the results of Angeli, De Leenheer, and Sontag do not apply.

Example 9.3. The following network is strongly endotactic, so it is permanent:

$$2A \rightleftharpoons B \qquad 2B \rightleftharpoons C \qquad 2C \rightleftharpoons A.$$

Combinatorially, the reactant polytope is a triangular prism; one triangle is formed by A, B, and C, and the other is formed by $2A$, $2B$, and $2C$. Each of the three pairs of reversible reactions lies along a diagonal of one of the quadrilateral boundary faces of the polytope.

Persistence in Example 9.3 already was known by work of Johnston and Siegel [29], which extends easily to the setting of mass-action differential inclusions, because the unique critical siphon $\{A, B, C\}$ can be shown to be weakly dynamically non-emptiable. In the next example, however, no previous results apply to obtain persistence: it is a three-dimensional network for which $\{A, B, C\}$ is a critical siphon but not weakly dynamically non-emptiable.

Example 9.4. The following network is strongly endotactic and therefore permanent:

$$4A \to A + B + C \to 4B \to 4C \to 2A + 2C.$$

The reactant polytope is a tetrahedron with vertices $A + B + C$, $4A$, $4B$, and $4C$. Each of the four reactions lies on or is a subset of one of the six edges of the tetrahedron.

9.2 Limitations of the geometric approach

The geometric approach presented in this work has some limitations in its ability to resolve Conjectures 4.3–4.7 in full generality. We explain these limitations via a simple example.

Example 9.5. The following network G is endotactic but not strongly endotactic:

$$0 \rightleftharpoons A \qquad B \rightleftharpoons 2B.$$

Define a toric jet adapted to G in the following way: take any sequence $(\theta(i))$ of numbers greater than 1 that limits to infinity, and let $w(i) = -(\frac{1}{\theta(i)}, 1)/\|(\frac{1}{\theta(i)}, 1)\|$. It is straightforward to verify that the conclusion of Proposition 6.26 fails for every subsequence of the

resulting toric jet $\left(\theta(i)^{w(i)}\right)$. Namely, the pull of the draining reaction $0 \leftarrow A$ with respect to the toric jet cannot be dominated by either of the two sustaining reactions $0 \rightarrow A$ or $B \rightarrow 2B$. Although results of Craciun, Nazarov, and Pantea [15, 39] can prove that this two-dimensional system is permanent, higher-dimensional generalizations of this network can not be proved to be permanent or even persistent.

The limiting direction of the unit jet here is $\lim_{n\to\infty} w(n) = (0, -1)$, and this direction $(0, -1)$ witnesses the fact that the network fails to be strongly endotactic. Namely, there is no reaction with reactant among the $\leq_{(0,-1)}$-maximal reactants $\{0, A\}$ and product outside of that set. Our results for strongly endotactic networks rely on the existence of such a sustaining reaction, so they are unable to prove permanence or even persistence of this network or similar networks by way of toric jets.

In our final example, the question of persistence is as yet unresolved.

Example 9.6. The following 3-dimensional network is weakly reversible—and therefore endotactic—but not strongly endotactic:

$$A \rightarrow B \rightarrow C \rightarrow A \qquad 2A \rightleftharpoons 3B.$$

The reactant polytope is a pyramid whose base has vertices A, $2A$, B, and $3B$. Three of the reactions form a triangle along edges of the pyramid, while the remaining pair of reversible reactions lies on the edge of the pyramid nonadjacent to the triangle. The network has a unique critical siphon, $\{A, B, C\}$, which is not weakly dynamically non-emptiable. Persistence and permanence (Conjectures 4.4 and 4.5) remain open for this network.

Acknowledgements. MG was supported by a Ramanujan fellowship from the Department of Science and Technology, India, and, during a semester-long stay at Duke University, by the Duke MathBio RTG grant NSF DMS-0943760. EM was supported by NSF grant DMS-1001437. AS was supported by an NSF postdoctoral fellowship DMS-1004380. The authors thank David F. Anderson, Gheorghe Craciun, Matthew Johnston, and Casian Pantea for helpful discussions, and Duke University, where many of the conversations occurred. The authors also thank two referees whose perceptive and insightful comments improved this work.

References

[1] Leonard Adleman, Manoj Gopalkrishnan, Ming-Deh Huang, Pablo Moisset, and Dustin Reishus, *On the mathematics of the law of mass action*, preprint 2008. `http://arXiv.org:0810.1108`

[2] David F. Anderson, *Boundedness of trajectories for weakly reversible, single linkage class reaction systems*, J. Math. Chem. **49** (2011), no. 10, 2275–2290.

[3] ———, *Global asymptotic stability for a class of nonlinear chemical equations*, SIAM J. Appl. Math. **68** (2008), no. 5, 1464–1476.

[4] ———, *A proof of the global attractor conjecture in the single linkage class case*, SIAM J. Appl. Math. **71** (2011), no. 4, 1487–1508.

[5] David F. Anderson and Anne Shiu, *The dynamics of weakly reversible population processes near facets*, SIAM J. Appl. Math. **70** (2010), 1840–1858.

[6] David Angeli, *On modularity and persistence of chemical reaction networks*, in *Decision and Control, 2008. CDC 2008. 47th IEEE Conference on*, pages 2650–2655, Dec. 2008.

[7] David Angeli, Patrick De Leenheer, and Eduardo Sontag, *Chemical networks with inflows and outflows: A positive linear differential inclusions approach*, Biotechnol. Progr. **25** (2009), 632–642.

[8] ———, *Graph-theoretic characterizations of monotonicity of chemical networks in reaction coordinates*, J. Math. Biol. **61** (2010), 581–616.

[9] ———, *Persistence results for chemical reaction networks with time-dependent kinetics and no global conservation laws*, SIAM J. Appl. Math. **71** (2011), 128–146.

[10] Murad Banaji and Janusz Mierczynski, *Global convergence in systems of differential equations arising from chemical reaction networks*, preprint 2012. `http://arXiv.org:1205.1716`

[11] M.W. Birch, *Maximum likelihood in three-way contingency tables*, J. Roy. Stat. Soc. B Met. **25** (1963), 220–233.

[12] A.D. Bruno. *Power Geometry in Algebraic and Differential Equations*. Elsevier, Amsterdam, 2000.

[13] Gheorghe Craciun, Alicia Dickenstein, Anne Shiu, and Bernd Sturmfels, *Toric dynamical systems*, J. Symb. Comput. **44** (2009), no. 11, 1551–1565.

[14] Gheorghe Craciun, Luis García-Puente, and Frank Sottile, *Some geometrical aspects of control points for toric patches*, Lect. Notes Comput. Sc., Revised selected papers of 7th International Conference on Mathematical Methods for Curves and Surfaces, **5862** (2010), 111–135.

[15] Gheorghe Craciun, Fedor Nazarov, and Casian Pantea, *Persistence and permanence of mass-action and power-law dynamical systems*, SIAM J. Appl. Math. **73** (2013), no. 1, 305–329.

[16] Jian Deng, Martin Feinberg, Chris Jones, and Adrian Nachman, *On the steady states of weakly reversible chemical reaction networks*, preprint, 2011. `http://arXiv.org:1111.2386`

[17] Pete Donnell and Murad Banaji. *Local and global stability of equilibria for a class of chemical reaction networks*, preprint, 2012. `http://arXiv.org:1211.2153`

[18] Martin Feinberg, *Chemical reaction network structure and the stability of complex isothermal reactors I. The deficiency zero and deficiency one theorems*, Chem. Eng. Sci. **42** (1987), no. 10, 2229–2268.

[19] ———, *Complex balancing in general kinetic systems*, Arch. Rational Mech. Anal. **49** (1972), no. 3, 187–194.

[20] Karin Gatermann, *Counting stable solutions of sparse polynomial systems*, Proceedings of an AMS-IMS-SIAM Joint Summer Research Conference on Symbolic Computation: Solving Equations in Algebra, Geometry, and Engineering, Mount Holyoke College, South Hadley, MA, June 11-15, 2000. Vol. 286. American Mathematical Soc., 2001.

[21] Gilles Gnacadja, *Univalent positive polynomial maps and the equilibrium state of chemical networks of reversible binding reactions*, Adv. in Appl. Math. **43** (2009), no. 4, 394–414.

[22] ———, *Reachability, persistence, and constructive chemical reaction networks (part III): a mathematical formalism for binary enzymatic networks and application to persistence*, J. Math. Chem. **49** (2011), no. 10, 2158–2176.

[23] Manoj Gopalkrishnan, *Catalysis in reaction networks*, B. Math. Biol. **73** (2011), no. 12, 2962–2982.

[24] Manoj Gopalkrishnan, Ezra Miller, Anne Shiu. *A projection argument for differential inclusions, with application to mass-action kinetics*, SIGMA **9** (2013), 025, 25 pages.

[25] F. Horn, *The dynamics of open reaction systems*, in *Mathematical aspects of chemical and biochemical problems and quantum chemistry, SIAM–AMS Proceedings, VIII*, pages 125–137, 1974.

[26] ———, *Necessary and sufficient conditions for complex balancing in chemical kinetics*, Arch. Ration. Mech. Anal. **49** (1972), no. 3, 172–186.

[27] Friedrich J. M. Horn and Roy Jackson, *General mass action kinetics*, Arch. Ration. Mech. Anal. **49** (1972), 81–116.

[28] Matthew D. Johnston and David Siegel, *Linear conjugacy of chemical reaction networks*, J. Math. Chem. **49** (2011), no. 7, 1263–1282.

[29] ———, *Weak dynamic non-emptiability and persistence of chemical kinetic systems*, SIAM J. Appl. Math. **71** (2011), no. 4, 1263–1279.

[30] Matthew D. Johnston, David Siegel, and Gábor Szederkényi, *Dynamical equivalence and linear conjugacy of chemical reaction networks: new results and methods*, MATCH–Commun. Math. Comput. Chem. **68** (2012), no. 2, 443–468.

[31] A. J. Lotka, *Analytical Note on Certain Rhythmic Relations in Organic Systems*, Proc. Natl. Acad. Sci. U.S. **6** (1920), 410–415.

[32] Sean Meyn and Richard L. Tweedie, *Markov chains and stochastic stability*, second edition. Cambridge University Press, 2009.

[33] Ezra Miller, *Theory and applications of lattice point methods for binomial ideals*, in Combinatorial Aspects of Commutative Algebra and Algebraic Geometry, Proceedings of the Abel Symposium (Voss, Norway, 1–4 June 2009), Abel Symposia, vol. 6, Springer Berlin Heidelberg, 2011, pp. 99–154.

[34] Ezra Miller and Igor Pak, *Metric combinatorics of convex polyhedra: cut loci and nonoverlapping unfoldings,* Discrete Comput. Geom. **39** (2008), no. 1–3, 339–388.

[35] Stefan Müller and Georg Regensburger, *Generalized mass action systems: Complex balancing equilibria and sign vectors of the stoichiometric and kinetic-order subspaces*, SIAM J. Appl. Math. **72** (2012), no. 6, 1926–1947.

[36] James Munkres. *Topology*. Prentice Hall, 2000.

[37] Lior Pachter and Bernd Sturmfels. *Algebraic statistics for computational biology*. Cambridge University Press, 2005.

[38] James R. Norris. *Markov chains*. Cambridge University Press, Cambridge series in statistical and probabilistic mathematics, 1998.

[39] Casian Pantea, *On the persistence and global stability of mass-action systems*, SIAM J. Math. Anal. **44** (2012), no. 3, 1636–1673.

[40] Carl Adam Petri, *Communication with automata*, Ph.D. dissertation, Technischen Hochschule Darmstadt (1962).

[41] Yuri Rabinovich, Alistair Sinclair, and Avi Wigderson, *Quadratic dynamical systems*, Proceedings of the 33rd Annual Symposium on Foundations of Computer Science (1992), 304–313.

[42] David Richeson and Jim Wiseman, *A fixed point theorem for bounded dynamical systems*, Illinois J. Math. **46** (2002), no. 2, 491–495.

[43] Michael A. Savageau and Robert Rosen, *Biochemical systems analysis: a study of function and design in molecular biology.* Addison-Wesley, Reading, MA, 1976.

[44] Guy Shinar and Martin Feinberg, *Concordant chemical reaction networks*, Math Biosci. **240** (2012), no. 2, 92–113.

[45] Anne Shiu and Bernd Sturmfels, *Siphons in chemical reaction networks*, B. Math. Biol. **72** (2010), no. 6, 1448–1463.

[46] David Siegel and Matthew D. Johnston, *A stratum approach to global stability of complex balanced systems*, Dyn. Syst. **26** (2011), 125–146.

[47] David Siegel and Debbie MacLean, *Global stability of complex balanced mechanisms*, J. Math. Chem. **27** (2000), no. 1–2, 89–110.

[48] Roman Srzednicki, *On rest points of dynamical systems*, Fund. Math. **126** (1985), no. 1, 69–81.

[49] Gábor Szederkényi and Katalin M. Hangos, *Finding complex balanced and detailed balanced realizations of chemical reaction networks*, J. Math. Chem. **49** (2011), no. 6, 1163–1179.

[50] J. Wei, *Axiomatic treatment of chemical reaction systems*, J. Chem. Phys. **36** (1962), 1578–1584.

[51] Günter Ziegler. *Lectures on Polytopes.* Vol. 152 of Graduate Texts in Mathematics. Springer–Verlag, New York, 1995.