KØBENHAVNS UNIVERSITET

Bachelorprojekt

Valdemar Skou Knudsen

Qualitative Aspects of Mathematical Reaction Networks Using Methods from Algebra and Degree Theory

Vejleder: Elisenda Feliu

Institut for Matematiske Fag. January 13, 2022

Abstract

The goal of this project is to give an introduction some areas within chemical reaction network theory. The structure of reaction networks and its interplay with the associated systems of differential equations is described, and different methods using tools from both algebra and analysis are introduced and applied to several concrete networks. A large part of the project is dedicated to a given network's capacity for multistationarity, which is an interesting property concerning the number of steady states for the associated ODE system. The method developed here involves computing the $\mathcal{C}^1$-mapping degree for a certain family of functions $\varphi_c \in \mathcal{C}^1(\mathbb{R}^n_{\geq 0}, \mathbb{R}^n)$. Since many of the properties of the mapping degree are nontrivial, their proofs are given in full detail. Finally, the project is wrapped up by applying all of the methods developed on several specific reaction networks.

Contents

1 Prerequisites

This project is written with a prior knowledge of commutative and homological algebra, algebraic geometry, all corresponding to a standard first course in the subject. Additionally, knowledge from standard bachelor-level courses in mathematics is also assumed.

2 Acknowledgements

I would like to thank my supervisor Elisenda Feliu for her insight and support throughout the last few months, and for her patience with me constantly changing the topic of the project.

I would also like to thank my friends and family for always showing interest in the project, and for even being willing to read some of the initial unpolished drafts. I took all of their comments to heart.

3 Reaction Networks and Systems

3.1 Basic Constructions and Examples

We begin by working through the basic notions of reaction networks and their associated systems of differential equations. The definition of a reaction network is best illustrated through an example:

Example 1. Consider the simple model of two basal protein productions and degradation, from gene X_1, X_2 respectively:

$$X_1 \longrightarrow X_1 + P_1, \quad X_2 \longrightarrow X_2 + P_2, \quad P_1 \longrightarrow 0, \quad P_2 \longrightarrow 0.$$

The model consists of four separate reactions, and the representation is meant to indicate that for each time one of the reaction occurs, the species written on the left-hand side is consumed (in the amount specified) and transformed into the species written on the right-hand side (in the amount specified). This particular model will occur as part of larger networks throughout several examples that will be seen later on. The model fits into the following framework, which is our first and main definition:

Definition 1. Let $\mathcal{X}$ be a finite set. A *chemical reaction network* (CRN) over $\mathcal{X}$ is a finite digraph $G = (\mathcal{Y}, \mathcal{R})$, where the set of nodes $\mathcal{Y}$ consists of formal $\mathbb{N}$-linear combinations

$$y_i = \sum_{i=k}^{n} y_{ik} X_k$$

of the elements of $\mathcal{X}$. We have that

- An element $X_i \in \mathcal{X}$ is called a *species*; we denote the cardinality $\#\mathcal{X} = n$.
- An element $y_i = (y_{i1}, \ldots, y_{in}) \in \mathcal{Y}$ is called a *complex*; we denote $\#\mathcal{Y} = m$.
- An element $y_i \to y_j \in \mathcal{R}$ is called a *reaction*; we denote $\#\mathcal{R} = r$.

Our previous example is just one of many biochemical models that fit into the structure of the above definition. The motivation for using such a framework is however not immediately obvious. The benefit is that we can associate several invariants to a given network, which might produce useful information.

Definition 2. Let G be a CRN over $\mathcal{X}$. The *stoichiometric subspace* S of G is the real vector-space generated by the vectors associated with the reactions $y_i \to y_j \in \mathcal{R}$:

$$S := \langle y_j - y_i : y_i \to y_j \in \mathcal{R} \rangle \subseteq \mathbb{R}^n.$$

Similarly, if $R_k \in \mathcal{R}$ is the reaction $y_i \to y_j$ then the k'th column of the *stoichiometric matrix* $N \in \mathbb{R}^{n \times r}$ is given by the vector $y_j - y_i$. We always write $s = \dim_{\mathbb{R}}(S) = \operatorname{rank}(N)$.

Note that the stoichiometric subspace only depends on the network structure. We now introduce the main object of study for this project, namely an associated system of differential equations to a given network.

Definition 3. Let G be a CRN over $\mathcal{X}$. A *mass-action network* arises from a labelling $(\mathcal{Y}, \mathcal{R}, (k_{ij})_{y_i \to y_j \in \mathcal{R}})$ of the directed edges of the network by real positive reaction-rate constants $k_{ij} \in \mathbb{R}_{>0}$. We then have an associated mass-action system $\dot{x} = f(x)$ defined by

$$\dot{x} = f(x) = \sum_{y_i \to y_j \in \mathcal{R}} k_{ij} x^{y_i} (y_j - y_i) = \sum_{y_i \to y_j \in \mathcal{R}} k_{ij} x_1^{y_{i1}} \cdots x_n^{y_{1n}} (y_j - y_i)$$

where $x_1, \ldots x_n$ are functions of t, and are thought of as concentrations of the species $X_1, \ldots, X_n$. The function f is called the *species-formation function.*

Note that these systems are autonomous, and that the mass-action assumption ensures that the species-formation function $f \in \mathcal{C}^1(\mathbb{R}^n_{\geq 0}, \mathbb{R}^n)$ consists of polynomials. Finding explicit nontrivial solutions to a mass-action, even for very simple networks, is very difficult. Additionally, The reaction-rate parameters are often unknown (in particular if one wishes to study the general dynamics that arise from the network), which means that numerical analysis will also often not be of much use. Most of the results that we consider will therefore be of a more qualitative type. From now on we will freely assume that every CRN G is also a mass-action network.

Definition 4. Let G be a CRN and $\dot{x} = f(x)$ the associated mass-action system. We say that $\lambda \in \mathbb{R}^n$ is a *conservation law* for f if $\langle \lambda, f(x) \rangle = 0$ for all $x \in \mathbb{R}^n$. The set of conservation laws is denoted Λ, and we say that $\dot{x} = f(x)$ is conservative if there exists conservation law $\lambda \in \Lambda \cap \mathbb{R}^n_{>0}$ with positive entries.

Note that the range of the formation rate function f is by definition contained in the stoichiometric subspace S, which means that $S^\perp \subseteq \Lambda$ is an inclusion of vector spaces. If $S^\perp \cap \mathbb{R}^n_{>0} \neq \emptyset$ then we say that the underlying network G is conservative, which clearly implies that the associated mass-action network is also conservative.

Example 2. Consider the mass-action network

$$E + S \underset{k_2}{\overset{k_1}{\rightleftharpoons}} Y_1, \longrightarrow E + P, \qquad F + P \underset{k_4}{\overset{k_3}{\rightleftharpoons}} Y_2 \longrightarrow F + S.$$

The associated stoichiometric matrix and subspace are given by

$$N = \begin{pmatrix} 1 & -1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & -1 \\ -1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 1 & 1 \\ 0 & 0 & 1 & -1 & 1 & 0 \\ -1 & 1 & 0 & 0 & 0 & 1 \end{pmatrix}, \qquad S = \operatorname{span}\left(\begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ -1 \\ -1 \\ 0 \end{pmatrix}\right).$$

A simple computation shows that

$$S^\perp = \operatorname{span}\left(\begin{pmatrix} 1 \\ 0 \\ 0 \\ -1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \\ -1 \\ -1 \end{pmatrix}\right),$$

and from this we deduce that it is indeed possible to construct a positive conservation law

$$2\begin{pmatrix} 1 \\ 0 \\ 0 \\ -1 \\ 1 \\ 1 \end{pmatrix} + 2\begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \\ -1 \\ -1 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix},$$

ergo the network is conservative.

Definition 5. Let G be a CRN and $\dot{x} = f(x)$ the associated mass-action system. A subset $P \subseteq \mathbb{R}^n$ is called *forward invariant* if, for any interval I with $0 \in I$ and trajectory $x \colon I \to \mathbb{R}^n$ of the system with $x(0) \in P$, it holds that $x(t) \in P$ for all $t \in I \cap \mathbb{R}_{\geq 0}$.

Let $x \colon I \to \mathbb{R}^n$ be a trajectory of the a mass action system $\dot{x} = f(x)$, I is some interval with $0 \in I$. Note that for any $\lambda \in \Lambda$ it holds that $\langle \lambda, x(t) \rangle = K_\lambda$ is constant for all $t \in I \cap \mathbb{R}_{\geq 0}$. In particular we see that

$$\langle x(t) - x(0), \lambda \rangle = \langle x(t), \lambda \rangle - \langle x(0), \lambda \rangle = 0,$$

which means exactly that $x(t) \in x(0) + \Lambda^\perp$, so the set is forward invariant. The exact same argument also shows that the trajectory $x \colon I \to \mathbb{R}^n$ lies in the linear variety $x(0) + S^\perp$.

Proposition 1. *The positive orthant $\mathbb{R}^n_{\geq 0}$ is forward invariant.*

The proof is omitted, but can be found in [9]. Forward invariance is preserved under set intersection, which motivates the following definition.

Definition 6. Let G be a CRN with stoichiometric subspace S. The *stoichiometric compatibility class* (often just called an S-class) of $x_0 \in \mathbb{R}^n_{\geq 0}$ is given by $(x_0 + S) \cap \mathbb{R}^n_{\geq 0}$.

Note that the set $(x_0 + S) \subseteq \mathbb{R}^n$ is an affine subspace, equivalently a linear variety. The S-class is therefore the restriction of this affine subspace to $\mathbb{R}^n_{\geq 0}$, often just called a nonnegative variety. From the previous arguments it is clear that every trajectory $x\colon I \to \mathbb{R}^n$ with $x(0) \in \mathbb{R}^n_{\geq 0}$ lies in some S-class. This somewhat unexpected linearity is a very useful property of these mass-action systems. We have the following:

Proposition 2. *Let G be a CRN and $\dot{x} = f(x)$ the associated mass-action system. The system is conservative if and only if $(x_0 + \Lambda^\perp) \cap \mathbb{R}^n_{\geq 0}$ is bounded for all $x_0 \in \mathbb{R}^n_{\geq 0}$. Likewise, the underlying network G is conservative if and only if $(x_0+S) \cap \mathbb{R}^n_{\geq 0}$ is bounded (and thus compact) for all $x_0 \in \mathbb{R}^n_{\geq 0}$.*

Proof. We only prove the forward implication in both cases, since this is the statement that we will frequently use. The reverse implication can be found in Theorem 4 in [2]. Additionally, since the proofs of statement is identical in both cases, we focus on the case where the system is conservative, since this is the more general statement. So, suppose that $\lambda \in \Lambda^\perp \cap \mathbb{R}^n_{>0}$ is a positive conservation law and let $m = \min(\lambda_1, \ldots, \lambda_n)$. Let $x \in \mathbb{R}^n_{\geq 0}$ and consider the nonnegative linear variety $(x + \Lambda) \cap \mathbb{R}^n_{\geq 0}$. For all $y \in (x + \Lambda) \cap \mathbb{R}^n_{\geq 0}$ we have that $\langle y, \lambda\rangle = \langle x, \lambda\rangle$, and therefore also that

$$y_k \lambda_k \leq \langle y, \lambda\rangle = \langle x, \lambda\rangle \quad \text{for all } 1 \leq k \leq n.$$

We therefore conclude that

$$0 \leq y_k \leq \frac{\langle x, \lambda\rangle}{\lambda_k} \leq \frac{\langle x, \lambda\rangle}{m}.$$

We therefore have a uniform bound for each of the coordinates of y, and since both y and x were arbitrary we conclude that $(x + \Lambda) \cap \mathbb{R}^n_{\geq 0}$ is bounded (and thus compact) for all $x \in \mathbb{R}^n_{\geq 0}$. □

Let us say a few more words about the structure of S-classes. We have that

$$d = \dim_{\mathbb{R}} S^\perp = \dim_{\mathbb{R}} \mathbb{R}^n - \dim_{\mathbb{R}} S = n - s,$$

so we can find a matrix $W \in \mathbb{R}^{d\times n}$ whose rows span $S^\perp$. Note that we can freely assume that W is row-reduced. Let $x_0 \in \mathbb{R}^n_{\geq 0}$ be arbitrary and note that

$$y \in (x_0 + S) \cap \mathbb{R}^n_{\geq 0} \Leftrightarrow W(y - x_0) = 0 \Leftrightarrow y \in \{x \in \mathbb{R}^n_{\geq 0} \mid Wx = Wx_0\}.$$

Notationwise we therefore, for : $c = Wx_0 \in \mathbb{R}^d$, denote the stoichiometric compatibility class of c by

$$\mathcal{P}_c = \{x \in \mathbb{R}^n_{\geq 0} \mid Wx = c\}.$$

In other words, the S-class is the level set at $c \in \mathbb{R}^d$ for the linear function given by the matrix W. We shall also use this notation for general matrices $W \in \mathbb{R}^{d\times n}$, i.e. in situations where we are not implicitly assuming that W is associated to some reaction network.

A trivially equivalent condition for the network being conservative is therefore that there exists a linear combination of the rows of W with each entry being positive.

Example 3. Consider the mass-action network

$$X_1 \underset{k_2}{\overset{k_1}{\rightleftharpoons}} 2X_1, \quad X_2 \underset{k_4}{\overset{k_3}{\rightleftharpoons}} X_1 + X_2 \underset{k_6}{\overset{k_5}{\rightleftharpoons}} X_3.$$

Similarly to the previous example, we have that

$$W = \begin{pmatrix} 0 & 1 & 1 \end{pmatrix} \in \mathbb{R}^{1\times 3}$$

is a row-reduced matrix whose rows generate $S^\perp$, where S is the stoichiometric subspace of the network. In this case, we see that the network is clearly not conservative.

3.2 Dissipative Networks

In many cases it does not hold that the network is conservative, i.e. not all of the stoichiometric compatibility classes $\mathcal{P}_c$ are bounded. For instance, all networks with an inflow $0 \to X_k$ or an outflow $X_k \to 0$ for some $X_k \in \mathcal{X}$ will not be conservative (this is seen through a routine computation). However, it may happen that all trajectories lie in some further compact subset $K_c \subseteq \mathcal{P}_c$. This is a special case of the upcoming definition, and we shall see that many of the nice properties of the network are preserved in this more general setting.

Definition 7. Let G be a CRN with associated mass-action system $\dot{x} = f(x)$. We say that G is *dissipative* if for all stoichiometric compatibility classes $\mathcal{P}_c$ with $\mathcal{P}_c^+ \neq \emptyset$, there exists a compact set $K_c \subseteq \mathcal{P}_c$ such that for all trajectories $x\colon I \to \mathbb{R}^n$ where $0 \in I$ is some interval, and $x(0) \in \mathcal{P}_c$ it holds that $x(t) \in K_c$ for all $t \geq t(x)$ for some $t(x) \geq 0$.

In other words, the network is dissipative if all trajectories $x\colon I \to \mathbb{R}^n$ with $x(0) \in \mathcal{P}_c$ are eventually uniformly bounded. This property means that all trajectories are defined for all $t \geq 0$, guaranteeing the existence of the global semiflow $\Phi(x,t)$ for the system. As expected, conservative networks are in particular dissipative, since we can simply pick $\mathcal{P}_c$ as the desired compact set. The set K_c is called attracting, and without any further assumptions can be chosen to have a number of desirable properties:

Lemma 1. *Suppose that G is a dissipative network, consider an S-class $\mathcal{P}_c$ such that $\mathcal{P}_c^+ \neq \emptyset$, and let $K_c \subseteq \mathcal{P}_c$ be the attracting set. There exists an attracting set $K_c' \supseteq K_c$ satisfying the following properties:*

- *$K_c' \cap \mathbb{R}^n_{>0} \neq \emptyset$.*
- *All ω-limit points of the system that are in $\mathcal{P}_c$ are also in K_c'.*
- *K_c' is forward invariant and all ω-limit points in $\mathbb{R}^n_{>0}$ are in the interior $K_c'^{\circ}$.*

Proof. For the first part of the proof, pick some $y \in \mathcal{P}_c^+$ (this is possible by assumption) and replace K_c with $K_c \cup \{y\}$ if necessary. This set is still attracting and compact (since compactness in metric spaces is preserved under union, and both K_c and $\{y\}$ are compact). Furthermore, it is clear that $(K_c \cup \{y\}) \cap \mathbb{R}^n_{>0} \neq \emptyset$, which was the desired property.

For the second part suppose for the sake of contradiction that there is some limit point $x' \in \mathcal{P}_c \setminus K_c'$. This means that there is a trajectory $x\colon I \to \mathbb{R}^n$ with $0 \in I$ and a sequence $(t_i)_{i\in\mathbb{N}}$ with $t_i \geq 0$ and $t_i \to \infty$ such that $x(t_i) \to x'$. Let $t(x)$ be the maximum entry time for the trajectory. Since K_c' is closed, there exists $\varepsilon > 0$ and $N \in \mathbb{N}$ such that $B_\varepsilon(x') \cap K_c' = \emptyset$ and such that $x(t_i) \in B_\varepsilon(x')$ for all $i \geq N$. However, we can freely pick $N \in \mathbb{N}$ sufficiently large such that $t_i \geq t(x)$ for all $i \geq N$, which is clearly a contradiction.

By potentially picking a slightly larger set, we can freely assume that K_c' contains all ω-limit points, and such that all ω-limit points that are not in $\partial\mathbb{R}^n_{\geq 0}$ are in the interior $K_c'^{\circ}$. Consider the function

$$\begin{aligned}\tau\colon \mathbb{R}^n_{\geq 0} &\to \mathbb{R}_{\geq 0}\\ x &\mapsto \inf\{t \geq 0\colon x(t) \in K_c'\},\end{aligned}$$

where $x\colon I \to \mathbb{R}^n_{\geq 0}$ is the unique solution corresponding to the initial condition $x(0) = x$. The value $\tau(x)$ is the first entry time of x into K_c'. Now, τ is locally bounded. Indeed, if this was not the case in some point $x_0 \in \mathbb{R}^n_{\geq 0}$ it would hold that

$$\forall \varepsilon > 0\, \exists (x_k)_{k\in\mathbb{N}} \subseteq B_\varepsilon(x_0)\colon \tau(x_k) \to \infty.$$

Letting $\varepsilon \to 0$ and using continuity of trajectories $x\colon I \to \mathbb{R}^n$ we get that $\tau(x_0) = \infty$, which is clearly a contradiction, since all ω-limit points are assumed to be in K_c', so we conclude that τ is indeed locally bounded. Now, we have an open covering

$$K_c' \subseteq \bigcup_{x\in K_c'} B_{\varepsilon(x)}(x),$$

where $\varepsilon(x) > 0$ is chosen such that τ is bounded in $B_{\varepsilon(x)}(x)$. By compactness, this can be reduced to a finite open subcovering, indexed by points $\{x_1, \dots, x_m\} \subseteq K'_c$ which means that

$$T = \max\{\tau(x)\colon x \in K'_c\} \leq \max\{\tau(x)\colon x \in B_{\varepsilon(x_1)}(x_1) \cup \cdots \cup B_{\varepsilon(x_m)}(x_m)\} < \infty,$$

where the finiteness follows from the fact that that τ is bounded on each set in the above union. Note that T is the maximum time for trajectories leaving K_c to re-enter K_c. In particular, we note that for all trajectories $x\colon I \to \mathbb{R}^n$ we have that $x(t) \in K_c$ for all $t \geq T$. We can then redefine K'_c to be the slightly larger set

$$K'_c = \{x(t)\colon x(0) \in K_c, 0 \leq t \leq T\} = \{x(t)\colon x(0) \in K_x, 0 \leq t\},$$

which is clearly compact and forward invariant. □

Verifying whether a given reaction network is dissipative is nontrivial. It turns out that a desired attracting set K_c can be constructed given the existence of a linear Lyapunov function.

Proposition 3. *Suppose that for all $c \in \mathbb{R}^d$ such that $\mathcal{P}_c^+ \neq \emptyset$ there exists $w \in \mathbb{R}^n_{>0}$ and $r > 0$ such that $w \cdot f(x) < 0$ for all $x \in \mathcal{P}_c$ with $\|x\| \geq r$. Then the network is dissipative.*

Proof. Let $c \in \mathbb{R}^d$ be such that $\mathcal{P}_c^+ \neq \emptyset$ and let $w \in \mathbb{R}^n_{>0}$ be given as above. Consider the linear function

$$V\colon \mathbb{R}^n_{\geq 0} \to \mathbb{R}^n$$
$$x \mapsto \langle w, x\rangle = \sum_{i=1}^n w_i x_i$$

Note that $V(0) = 0$ and $V(x) \geq 0$ for all $x \in \mathbb{R}^n_{\geq 0}$. Furthermore, a restatement of our initial assumption is exactly that for all trajectories $\Phi(x,t)$ in $\mathcal{P}_c$ with $\|\Phi(x,t)\|$ we have that

$$\dot{V}(\Phi(x,t)) = \langle w, f(x)\rangle < 0.$$

Now, this means we can pick $R > 0$ such that

$$\{x \in \mathbb{R}^n_{\geq 0}\colon \ \|x\| \leq r\} \subseteq \{x \in \mathbb{R}^n_{\geq 0}\colon V(x) \leq R\},$$

and we construct the set

$$K_c := \{x \in \mathbb{R}^n_{\geq 0}\colon V(x) \leq R\} \cap \mathcal{P}_c.$$

We wish to show that this set satisfies the desired properties in order for the network to be dissipative. Now, suppose for the sake of contradiction that K_c is not attracting, i.e. that there exists $x' \in \mathcal{P}_c \setminus K_c$ such that $V(\Phi(x',t))$ is decreasing for all $t \geq 0$ and bounded below by R. In particular this means that

$$\lim_{t\to\infty} V(\Phi(x',t)) = R' \geq 0.$$

This means that for all $\varepsilon > 0$ we have that $\Phi(x',t) \in B_\varepsilon := \{x\colon V(x) \leq R' + \varepsilon\}$ for $t > 0$ large enough. B_ε is compact, so it must contain at least one ω-limit point x'' of $\Phi(x',t)$. This means that $V(x'') = R'$, and since the set of ω-limit points is forward invariant it must additionally hold that $\dot{V}(x'') = 0$. Since $\|x''\| \geq r$, this contradicts our initial assumptions. Therefore, we must have $\Phi(x',t) \in K_c$ for some $t(x') > 0$, so K_c is attracting, which means exactly that the network is dissipative. □

Importantly, the converse to the above proposition does not hold. However, for many concrete networks that model biochemical processes it is possible to construct find a suitable $w \in \mathbb{R}^n_{>0}$.

Example 1, continued. We now return to our first motivating example. Assume that the network has mass-action kinetics. The stoichiometric subspace is given by

$$S = \operatorname{span}\left(\begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right),$$

and a simple computation therefore yields that

$$W = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix},$$

is a row-reduced matrix whose rows generate $S^{\perp}$. We observe that the generating conservation laws of the network are $(1,0,0,0)^{\perp}$ and $(0,1,0,0)^{\perp}$; so the two species X_1, X_2 must be of constant concentration K_1, K_2 respectively. We also observe that this network is not conservative, since there is no linear combination of the generating conservation laws that has strictly positive entries. The species-formation function is given by

$$f(x) = \begin{pmatrix} 0 \\ 0 \\ k_1 x_1 - k_3 p_2 \\ k_2 x_2 - k_4 p_2 \end{pmatrix}.$$

For any $w \in \mathbb{R}^4_{>0}$ it therefore holds that

$$w \cdot f(x) < 0 \Leftrightarrow w_3 k_1 K_1 + w_4 k_2 K_2 < w_3 k_3 p_1 + w_4 k_4 p_2,$$

which is clearly true for all $p_1, p_2 > 0$ large enough, so from Theorem 3 we conclude that the network is indeed dissipative.

3.3 Steady States and Multistationarity

Definition 8. Let G be a CRN and $\dot{x} = f(x)$ the associated mass-action system. The *steady state variety* is defined as

$$V(f) = V(f_1, \ldots, f_n) = \{x \in \mathbb{R}^n_{\geq 0} \colon f_1(x) = \ldots = f_n(x) = 0\}.$$

Any element $x \in V(f)$ is called a *steady state*, and if $x_1, \ldots, x_n > 0$ we say that x is a positive steady state. The *positive steady state variety* is defined as $V_{>0}(f) := V(f) \cap \mathbb{R}^n_{>0}$.

Note that, unlike in most courses in algebraic geometry, we are only interested in solutions from the positive orthant of $\mathbb{R}^n$. This means that, if we want to study the steady states of a mass-action system, we are not immediately able to use tools from classical algebraic geometry. We are often interested in determining whether a given system has multiple steady states for the same set of parameters, and thus introduce the following notion:

Definition 9. Let G be a CRN. The associated mass-action system $\dot{x} = f(x)$ (for fixed parameters) is said to exhibit *multistationarity* if

$$\# V_{>0}(f) \cap \mathcal{P}_c \geq 2,$$

for some $c \in \mathbb{R}^d$. We say that G has the capacity for multistationarity if there exists a set of parameters such that the mass-action system exhibits multistationarity.

Note that the idea of looking at $V_{>0}(f) \cap \mathcal{P}_c$ is the correct way to analyse the number of steady states the system. This is because we have previously shown that any trajectory will be fully contained in the S-class $\mathcal{P}_c$ for some $c \in \mathbb{R}^d$. Simply finding all the solutions to $f(x) = 0$ will therefore not suffice. In general, this is therefore a much harder problem, but it becomes easier due to the (possibly very useful) conservation laws of the network as well as other properties that we might have deduced. In order to systematically determine whether a a system exhibits multistationarity we introduce an auxiliary construction:

Definition 10. Let $f \in \mathcal{C}^1(\mathbb{R}^n_{\geq 0}, \mathbb{R}^n)$ and let $W \in \mathbb{R}^{d \times n}$ be a row reduced matrix of maximal rank d, with indices $i_1, \ldots, i_d$ denoting the first non-zero coordinate of each row. For any $c \in \mathbb{R}^d$ the function $\varphi_c \in \mathcal{C}^1(\mathbb{R}^n_{\geq 0}, \mathbb{R}^n)$ constructed from f and W is defined by the coordinate functions

$$\varphi_c(x)_i = \begin{cases} f_i(x), & i \notin \{i_1, \ldots, i_d\}, \\ (Wx - c)_i & i \in \{i_1, \ldots, i_d\}. \end{cases}$$

We recall that the open and closed positive level sets of W at c are defined as

$$\mathcal{P}_c := \{x \in \mathbb{R}^n_{\geq 0} \mid Wx = c\} \qquad \mathcal{P}_c^+ := \{x \in \mathbb{R}^n_{>0} \mid Wx = c\}$$

By a suitable reordering of coordinates, we can assume that $\{i_1, \ldots, i_d\} = \{1, \ldots, d\}$ without changing the function φ_c. This means that we have an equality of block matrix

$$W = \begin{pmatrix} I_d & \widehat{W} \end{pmatrix},$$

where $\widehat{W} \in \mathbb{R}^{d \times (n-d)}$. Consider projection $\pi \colon \mathbb{R}^n \to \mathbb{R}^s$ onto the last $s := n - d$ coordinates, and note that by the above equality it holds for all $x, y \in \mathbb{R}^n$ with $Wx = Wy$ that $x = y$ if and only if $\pi(x) = \pi(y)$.

At first glance this construction might seem arbitrary, however it is motivated by the following special case. Let G be a CRN and $\dot{x} = f(x)$ the associated mass-action system. In this situation, we have a row-reduced matrix $W \in \mathbb{R}^d \times n$, whose rows span $S^\perp$ (where S is the stoichiometric subspace). By definition $Wf(x) = 0$ for all $x \in \mathbb{R}^n$, which means exactly that

$$x \in V_{>0}(f) \cap \mathcal{P}_c \Leftrightarrow x \in V_{>0}(\varphi_c).$$

The function φ_c thus records the exact data that we are interested in, which is very convenient. Later, when we return to studying this function, we will see that a lot of its properties are directly related to the structure of the underlying network.

4 Boundary Steady States for Networks

4.1 Siphons and Network Structure

So far, we have briefly considered criteria for whether a given mass-action system $\dot{x} = f(x)$ has bounded trajectories, and we mentioned a possible avenue for analysing the possibility for multistationarity. Recall that boundedness only depends on the network structure (namely on the existence of positive conservation laws).

We now wish to similarly study the existence of boundary steady states of the system $\dot{x} = f(x)$, which will include some more subtle aspects of the network structure.

Definition 11. Let G be a CRN and $\dot{x} = f(x)$ the associated mass-action system. We say that $x' \in \mathbb{R}^n_{\geq 0}$ is a boundary steady state if

$$x' \in \partial\mathbb{R}^n_{\geq 0} \cap V(f),$$

and we say that the network G has the capacity for boundary steady states if there exists a set of parameters such that the mass-action system has a boundary steady state.

In general, networks with the capacity for boundary steady states are somewhat difficult to work with for a number of reasons. In particular, as we shall see later, systematically studying the possibility for the system $\dot{x} = f(x)$ to exhibit multistationarity within some compatibility class $\mathcal{P}_c$ requires us to preclude the existence of steady states on the boundary $\partial\mathcal{P}_c$. It turns out that in many cases, the structure of the network alone will be sufficient to preclude its capacity for boundary steady states.

Definition 12. Let G be a CRN over $\mathcal{X} = \{X_1, \dots, X_n\}$. A nonempty set of species $Z \subseteq \mathcal{X}$ is called a siphon if for all $X_z \in Z$ and reactions $y_i \to y_j \in \mathcal{R}$ with $y_{jz} > 0$, then there exists $X_{z'} \in Z$ such that $y_{iz'} > 0$. Siphons that are minimal with respect to inclusion are called minimal siphons.

Intuitively, a species contained in a siphon is therefore only produced in a reaction if another species in that same siphon is also simultaneously consumed. With any siphon $Z \subseteq \mathcal{X}$ we consider the associated prime ideal $\mathfrak{p}_Z := \langle Z \rangle \subseteq \mathbb{Q}[X_1, \dots, X_n]$ generated by the set of species, as well as the corresponding set of indices viewed as a subset of $[n]$. The first connection between siphons and the existence of boundary steady states is given by the following lemma.

Lemma 2. *Let G be a a CRN over $\mathcal{X}$ and let $\dot{x} = f(x)$ be an associated mass-action system. If $\gamma \in \mathbb{R}^n_{\geq 0}$ is a boundary steady state, then the zero coordinate set $Z = \{X_i \in \mathcal{X} \colon \gamma_i = 0\}$ is a siphon.*

Proof. Let $X_z \in Z$ and let $I \subseteq \mathcal{R}$ be the set of reactions in which the reactant does not contain the species X_z but the product does:

$$I = \{y_i \to y_j \in \mathcal{R} \colon y_{iz} = 0 \text{ and } y_{jz} > 0\}.$$

It then follows from a simple calculation that

$$f_z(\gamma) = \sum_{y_i \to y_j \in I} k_{ij} y_{jz} \gamma^{y_i} = 0,$$

where the second equality follows because γ is a steady state. Since each term in the sum is non-negative, we must have $\gamma^{y_i} = \gamma_1^{y_{i1}} \cdots \gamma_n^{y_{in}} = 0$ for all $y_i \to y_j \in I$. Clearly, there must then exist a species $X_{z'} \in \mathcal{X}$ (which will depend on the reaction $y_i \to yj$) such that $\gamma_{z'}^{y_{iz'}} = 0$, but this is equivalent to $X_{z'} \in Z$ and $y_{iz'} > 0$, which means that Z is a siphon. $\square$

Consider an S-class $\mathcal{P}_c$ for some $c \in \mathbb{R}^d$, and a siphon (or also just some general index set) $Z \subseteq \mathcal{X}$. We let

$$F_{Z,c} := V_{\geq 0}(\mathfrak{p}_Z) \cap \mathcal{P}_c = \{x \in \mathcal{P}_c \colon x_i = 0 \text{ for all } i \in Z\} \subseteq \mathbb{R}^n$$

denote the unique face of $\mathcal{P}_c$ associated with Z. Note that for the zero coordinate sets of Lemma 2 we have that $\gamma \in F_{Z,c}$ by definition.

Definition 13. A siphon $Z \subseteq \mathcal{X}$ of G is called relevant if there exists a $c \in \mathbb{R}^d$ such that both $\mathcal{P}_c^+ \neq \emptyset$ and $F_{Z,c} \neq \emptyset$. If this is not the case we say that Z is nonrelevant.

A siphon being nonrelevant means that the face $F_{Z,c}$ cannot contain some steady state in any S-class $\mathcal{P}_c$ with $\mathcal{P}_c^+ \neq \emptyset$. In particular we have the following statement.

Proposition 4. *Let G be a CRN and suppose that all its siphons are nonrelevant. Then the associated mass-action system $\dot{x} = f(x)$ has no boundary steady states in any stoichiometric compatibility class $\mathcal{P}_c$ with $\mathcal{P}_c^+ \neq \emptyset$ (and for any choice of reaction-rate parameters).*

Proof. Suppose for the sake of contradiction that $\gamma \in \mathbb{R}^n_{\geq 0} \cap \mathcal{P}_c$ is a (boundary) steady state of the system, and that $\mathcal{P}_c^+ \neq \emptyset$. Then, it follows from Lemma 2 that the associated zero coordinate set $Z = \{X_i \in \mathcal{X} \colon \gamma_i = 0\}$ is a siphon. However then $\gamma \in F_{Z,c}$, which contradicts the assumption that Z in particular must be nonrelevant. □

Nonrelevant siphons are therefore desirable if we wish to preclude the existence of boundary steady states. We end with a criteria to determine whether a given siphon is nonrelevant.

Lemma 3. *A siphon $Z \subseteq \mathcal{X}$ of G is nonrelevant if and only if there exists a conservation law $l \in S^\perp \cap \mathbb{R}^n_{\geq 0}$ with positive non-negative entries satisfying supp$(l) = \{X_i \colon l_i > 0\} \subseteq Z$.*

Proof. We once again only prove the forward implication, since this is the statement that we will frequently use. The reverse implication can be found in Lemma 3.4 in [8]. So, let $l \in S^\perp \cap \mathbb{R}^n_{\geq 0}$ be a non-negative conservation relation with $\operatorname{supp}(l) \subseteq Z$, and suppose for the sake of contradiction that Z is relevant, i.e. that $F_{Z,c} \neq \emptyset$ and $\mathcal{P}_c^+ \neq \emptyset$ for some $c \in \mathbb{R}^d$. Pick two elements $x \in F_{Z,c}$ and $x' \in \mathcal{P}_c^+$. Since $\operatorname{supp}(x) \cap \operatorname{supp}(l) = \emptyset$ and $W(x - x') = c - c = 0$ (meaning that $x - x' \in S$ by definition of W) we have that

$$0 = \langle l, x - x' \rangle = \langle l, x \rangle - \langle l, x' \rangle = \langle l, x' \rangle$$

However, since both l and x' have non-negative entries we must have $x'_i = 0$ for all $i = 1, \ldots, n$ which is a contradiction. □

In general, determining all the siphons of a given network is quite difficult. Therefore, one instead computes the siphons that are minimal with respect to inclusion.

Proposition 5. *Let G be a CRN and suppose that all the siphons that are minimal with respect to inclusion are nonrelevant. Then all siphons of the network are nonrelevant.*

Proof. Suppose for the sake of contradiction that $Z \subseteq \mathcal{X}$ is a relevant siphon, i.e. that $F_{Z,c} \neq \emptyset$ for some $c \in \mathbb{R}^d$ with $\mathcal{P}_c^+ \neq \emptyset$. By assumption Z contains some minimal siphon Z'. However, by definition we have that

$$F_{Z,c} = V_{\geq 0}(\mathfrak{p}_Z) \cap \mathcal{P}_c \subseteq V_{\geq 0}(\mathfrak{p}_{Z'}) \cap \mathcal{P}_c = F_{Z',c},$$

which clearly contradicts the assumption that Z' is nonrelevant. □

The statement of Proposition 4 therefore also holds when considering only the minimal siphons of the network. Now, it is important to note that Proposition 4 is not a bi-implication. There exists networks with relevant minimal siphons for which it still is not possible to have any boundary steady states in any S-class $\mathcal{P}_c$ with $\mathcal{P}_c^+ \neq \emptyset$.

4.2 An Algebraic Characterization

We now proceed to give a characterization of the minimal siphons of a CRN G in terms of elementary ring theory, which was first introduced in [8]. This means that we have a systematic way of identifying and computing the siphons, which is extremely useful for working with specific reactions networks. Let us first introduce a certain way of classifying the structure of the network.

Definition 14. Let G be a CRN. We say that two complexes y_i and y_j are *connected* if there exists a sequence of reactions $y_i \to \cdots \to y_j$ or a sequence of reactions $y_j \to \cdots \to y_i$.

We can freely introduce reactions $X_i \to X_i$ for all $X_i \in \mathcal{X}$ without affecting the associated mass-action system of the network. Introducing the trivial reactions means that the connected complexes partition the network (i.e. the above definition is an equivalence relation between complexes). These equivalence classes are called the connected components of the network.

Definition 15. Let G be a CRN and G' be a connected component. If there for every $y_i, y_j \in G'$ exists both of the two sequences of reactions $y_i \to \cdots y_j$ and $y_j \to \cdots y_i$ then we say that the component is *strongly connected*. If the network G consist of one strongly connected component then we say that G is *strongly connected*.

Note that if G is strongly connected then for any pair $y_i, y_j \in G$ of complexes there exists a sequence of reactions $y_i \to \cdots \to y_j$. These connected components become important in the upcoming theorem. Let $R = \mathbb{Q}[X_1, \ldots, X_n]/\langle X_1 \cdots X_n \rangle$ denote the ring of polynomial functions with rational coefficients on the union of the coordinate hyperplanes in $\mathbb{R}^n$, and consider the following three ideals in R:

$$\begin{aligned}
\mathfrak{m}_{1,G} &:= \langle X^{y_i} \cdot (X^{y_j} - X^{y_i}) : y_i \to y_j \in \mathcal{R} \rangle, \\
\mathfrak{m}_{2,G} &:= \langle X^{y_j} - X^{y_j} : y_i \to y_j \in \mathcal{R} \rangle, \\
\mathfrak{m}_{3,G} &:= \langle X^{y_i} : y_i \in \mathcal{Y} \rangle.
\end{aligned}$$

Theorem 1. *The minimal siphons of G are equal to the inclusion-minimal sets $\{X_1, \ldots, X_n\} \cap \mathfrak{p}$ where $\mathfrak{p}$ runs over the minimal primes of $\mathfrak{m}_{1,G}$. If each connected component of G is strongly connected then $\mathfrak{m}_{1,G}$ can be replaced with $\mathfrak{m}_{2,G}$. If G is strongly connected then $\mathfrak{m}_{1,G}$ can be replaced with $\mathfrak{m}_{3,G}$.*

Proof. Consider the set

$$V_{\{0,1\}}(\mathfrak{m}_{1,G}) := \{x \in \{0,1\}^n \colon f(x) = 0 \text{ for all } f \in \mathfrak{m}\}.$$

So $V_{\{0,1\}}(\mathfrak{p}_{1,G})$ is the restriction of the variety $V(\mathfrak{m}_{1,G})$ to the set of binary n-tuples $\{0,1\}^n$. We first claim that the inclusion minimal sets $\{X_1, \ldots, X_n\} \cap \mathfrak{p}$, where $\mathfrak{p}$ runs over the minimal primes of $\mathfrak{m}_{1,G}$, can equivalently be written as minimal sets of the form $\operatorname{supp}(\gamma)^c = \{X_i \in \mathcal{X} \mid \gamma_i = 0\}$, where $\gamma \in V_{\{0,1\}}(\mathfrak{m}_{1,G})$. For this purpose, let $\mathfrak{p} \supset \mathfrak{m}_{1,G}$ be a minimal prime and define $\gamma \in \{0,1\}^n$ by

$$\gamma_i = \mathbf{1}_{\mathfrak{p}^c}(X_i) = \begin{cases} 1 & \text{if } X_i \notin \mathfrak{p} \\ 0 & \text{otherwise} \end{cases}$$

From this construction it clearly follows that

$$\{X_1, \ldots, X_n\} \cap \mathfrak{p} = \operatorname{supp}(\gamma)^c, \qquad\qquad \gamma \in V_{\{0,1\}}(\mathfrak{p}) \subset V_{\{0,1\}}(\mathfrak{m}_{1,G}),$$

where we have used that $\mathfrak{m}_{1,G} \subseteq \mathfrak{p}$ implies that $V_{\{0,1\}}(\mathfrak{p}) \subseteq V_{\{0,1\}}(\mathfrak{m}_{1,G})$. Conversely, suppose that $\gamma' \in V_{\{0,1\}}(\mathfrak{m}_{1,G})$ is chosen such that the set $\operatorname{supp}(\gamma')^c$ is minimal. From basic commutative algebra (see for instance Chapter 1 in [1]) we have that

$$\begin{aligned}
V_{\{0,1\}}(\mathfrak{m}_{1,G}) &= V_{\{0,1\}}\left(\sqrt{\mathfrak{m}_{1,G}}\right) \\
&= V_{\{0,1\}}\left(\bigcap_{\mathfrak{p} \supseteq \mathfrak{m}_{1,G}} \mathfrak{p}\right) \\
&= \bigcup_{\mathfrak{p} \supseteq \mathfrak{m}_{1,G}} V_{\{0,1\}}(\mathfrak{p}),
\end{aligned}$$

where $\sqrt{\mathfrak{m}_{1,G}}$ is the radical of the ideal $\mathfrak{m}_{1,G}$. In particular this means that $\gamma' \in V_{\{0,1\}}(\mathfrak{p}')$ for some minimal prime $\mathfrak{p}' \supseteq \mathfrak{m}_{1,G}$. The minimality of γ' ensures that $\operatorname{supp}(\gamma')^c = \{X_1, \ldots, X_n\} \cap \mathfrak{p}'$, which was what we wanted. We have therefore shown that the inclusion-minimal sets of the form $\{\mathfrak{p} \cap \{X_1, \ldots, X_n\} \colon \mathfrak{p} \supset \mathfrak{m}_{1,G} \text{ minimal prime}\}$ are exactly equal to the inclusion-minimal sets of the form $\{\mathrm{supp}(\delta)^c : \delta \in V_{\{0,1\}}(\mathfrak{m}_{1,G})\}$. However, by definition we have that all elements $\delta \in V_{\{0,1\}}(\mathfrak{m}_{1,G})\}$ satisfy that

$$\delta^{y_i} \cdot (\delta^{y_j} - \delta^{y_i}) = 0 \Longleftrightarrow \big(\delta^{y_i} \neq 0 \Rightarrow \delta^{y_j} \neq 0\big),$$

which is the defining characteristic for siphons, so we are done.

Now, suppose that the connected components of G are strongly connected, which means that for all reactions $y_i \to y_j \in \mathcal{R}$ there exists a sequence of reactions $y_j \to \cdots \to y_i$. Suppose first for the sake of simplicity that $y_j \to y_l, y_l \to y_i \in \mathcal{R}$. In particular this means that both

$$X^{y_j}(X^{y_l} - X^{y_j}, X^{y_l}(X^{y_i} - X^{y_l}) \in \mathfrak{m}_{1,G}$$

Working in $R/(\mathfrak{m}_{1,G})$ we see that

$$\begin{aligned} X^{y_j}(X^{y_i} - X^{y_j}) &= X^{y_i}X^{y_j} - X^{y_j}X^{y_j} \\ &= X^{y_j}X^{y_j} - X^{y_j}X^{y_l} \\ &= X^{y_j}(X^{y_j} - X^{y_l}) \\ &= 0, \end{aligned}$$

which means that ${}^{y_j}(X^{y_i} - X^{y_j}) \in \mathcal{R}$. Note that this argument easily extends via induction to the case where the reverse chain of reactions is longer than 2. Now, since also $X^{y_i}(X^{y_j} - X^{y_i}) \in \mathcal{R}$ by assumption, then so is the difference between these two elements

$$-(X^{y_i} - X^{y_j})^2 \in \mathcal{R}.$$

So, $(X^{y_i} - X^{y_j}) \in \sqrt{\mathfrak{m}_{1,G}}$, and since the reaction $y_i \to y_j$ was arbitrary we conclude that

$$\sqrt{\mathfrak{m}_{1,G}} \subseteq \sqrt{\mathfrak{m}_{2,G}},$$

and the reverse inclusion is true in general since $\mathfrak{m}_{2,G} \subseteq \mathfrak{m}_{1,G}$. Now, from commutative algebra we know that the prime ideals containing some ideal I only depend on its radical $\sqrt{I}$, so we therefore conclude that the two ideals $\mathfrak{m}_{1,G}$ and $\mathfrak{m}_{2,G}$ have the same minimal primes.

Lastly, suppose that G is strongly connected. In particular this means that the connected component(s) of G are strongly connected, so the ideal of interest is $\mathfrak{m}_{2,G}$. However, since G is strongly connected we know that for all pairs of complexes y_i, y_j both $X^{y_j} - X^{y_i} \in \mathfrak{m}_{2,G}$ and $X^{y_i} - X^{y_j} \in \mathfrak{m}_{2,G}$, which means that both $X^{y_i}, X^{y_j} \in \mathfrak{m}_{2,G}$. By definition we therefore have that $\mathfrak{m}_{2,G} = \mathfrak{m}_{3,G}$, and we are done. □

The minimal primes of a given ideal can be computed using standard algebra software. In the following examples, `Macaulay2` [6] is used for these computations, specifically the decompose command.

Example 1, continued. Recall that the generating conservation laws for this network was $x_1 = K_2$ $x_2 = K_2$. Since none of the components of the network are strongly connected it follows from Theorem 1 that the ideal of interest is given by

$$\mathfrak{m}_{1,G} = \langle X_1 \cdot (X_1X_3 - X_1), X_2 \cdot (X_2X_4 - X_2), X_3, X_4 \rangle.$$

Using `Macaulay2` we see that the unique minimal prime containing $\mathfrak{m}_{1,G}$ is $\langle X_1, X_2, X_3, X_4 \rangle$, which means that the unique minimal siphon of the network is just the full set of species $\mathcal{X} = \{X_1, X_2, X_3, X_4\}$. This is easily verified, and should also be obvious from the network structure. It therefore follows from a trivial application of Corollary 4 that the network does not contain any boundary steady states in any S-class $\mathcal{P}_c$ with $\mathcal{P}_c^+ \neq \emptyset$.

Example 4. Consider the following mass-action network, which is a T-cell signal transduction model.

$$\begin{tikzcd} T+M \arrow[r, "k_1"] & X_1 \arrow[r, "k_2"] \arrow[l, bend right, "k_3"'] & X_2 \arrow[ll, bend right=60, "k_4"'] \end{tikzcd}$$

Similarly to previous examples, we compute that

$$W = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & -1 \end{pmatrix}$$

is a row-reduced matrix whose rows generate $S^{\perp}$, where S is the stoichiometric subspace. The details are omitted. The generating conservation laws, and a positive conservation law (meaning that the network is conservative) are given by

$$\begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \qquad \begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \end{pmatrix}, \qquad 2\begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \\ 1 \\ 1 \end{pmatrix}.$$

Let us relabel $X_3 = T$ and $X_4 = M$. Since the network is strongly connected, it follows from the third part of Theorem 1 that the binomial ideal of interest is

$$\mathfrak{m}_{3,G} = \langle X_1, X_2, X_3X_4 \rangle.$$

Using `Macaulay2` we see that the minimal primes for this ideal are given by $\langle X_1, X_2, X_3 \rangle$ and $\langle X_1, X_2, X_4 \rangle$, which means that the minimal siphons of the network are given by $\{X_1, X_2, X_3\}$ and $\{X_1, X_2, X_4\}$. Both of these siphons are nonrelevant, since they contain the support of the non-negative conservation laws $(1, 1, 1, 0)^{\perp}$ and $(1, 1, 0, 1)^{\perp}$ respectively. Therefore it follows from Proposition 4 that the network does not contain any boundary steady states in any $\mathcal{P}_c$ with $\mathcal{P}_c^+ \neq \emptyset$.

5 Degree Theory and Regions of Multistationarity

5.1 The $\mathcal{C}^1$-Mapping Degree for Regular Values

Let G be a CRN. Our main goal is to give conditions that ensure whether G has the capacity for multistationarity or not. Additionally, we also wish to determine whether a given set of reaction-rate parameters will allow for the associated mass-action system $\dot{x} = f(x)$ to exhibit multistationarity. As mentioned in the previous section, the method will be to study the function φ_c as well as using the mapping degree for $\mathcal{C}^1$-functions. Since the degree and its properties are nontrivial, we first give quite a lengthy exposition to the relevant theory, roughly following [10]. A benefit of this is that we also prove some other interesting theorems along the way, and provide motivation for similar applications.

Recall that for a polynomial $f \in \mathbb{R}[x]$ of degree $d > 0$, the fundamental theorem of algebra ensures the existence of exactly d (possibly complex) roots. The goal of this new degree will be to provide information about the cardinality of the preimage $f^{-1}(p)$ for some arbitrary $f \in \mathcal{C}^1(\Omega, \mathbb{R}^n)$ and $p \in \mathbb{R}^n$, where $\Omega \subseteq \mathbb{R}^n$. So hopefully we can obtain results similar to the fundamental theorem of algebra, but for a much wider class of functions. We begin with a preliminary result, which will also be crucial for the constrution of the degree.

Definition 16. Let $f \in \mathcal{C}^1(\overline{\Omega}, \mathbb{R}^n)$ where $\Omega \subseteq \mathbb{R}^n$ is open and bounded, and $\overline{\Omega}$ is its closure. We say that $p \in \mathbb{R}^n$ is a regular value (for f) if the Jacobian $J_f(x)$ is invertible for all $x \in f^{-1}(p)$.

Lemma 4. *Suppose that $f \in \mathcal{C}^1(\overline{\Omega}, \mathbb{R}^n)$ where $\Omega \subseteq \mathbb{R}^n$ is open and bounded, and let $p \notin f(\partial\Omega)$ be a regular value. Then the preimage $f^{-1}(p)$ is a finite set.*

Proof. Suppose for the sake of contradiction that $f^{-1}(p)$ is infinite. From continuity we know that $f^{-1}(p)$ is also compact, which means that we can pick a limit point $x_0 \in f^{-1}(p)$ as well as a sequence $(x_k)_{k\in\mathbb{N}} \subseteq f^{-1}(p)$ such that $\|x_k - x_0\| \leq 1/k$. Taking a first order Taylor expansion of f around x_0 yields

$$f(x_0 + \xi) = f(x_0) + J_f(x_0)\xi + R_{x_0}(\xi),$$

where $\xi \in \mathbb{R}^n$. Since J_f is uniformly continuous on $\overline{\Omega}$ we have that

$$\forall \varepsilon > 0 \quad \exists \delta > 0 : \, t\,\|\xi\| < \delta \;\Rightarrow\; \|J_f(x_0 + t\xi) - J_f(x_0)\| < \varepsilon,$$

which allows for an estimate on the remainder term:

$$\begin{aligned} \|R_{x_0}(\xi)\| &= \left\| \int_{[0,1]} J_f(x_0 + t\xi) - f'(x_0)\xi \,\mathrm{d}t \right\| \\ &\leq \int_{[0,1]} \varepsilon \,\|\xi\| \,\mathrm{d}t \\ &= \varepsilon \,\|\xi\| \,. \end{aligned}$$

Since p is a regular value we know that $J_f(x_0)$ is invertible, and in particular there exists $c > 0$ such that $\|J_f(x_0)\xi\| \geq c\,\|\xi\|$ for all $\xi \in \mathbb{R}^n$. Now, suppose that $\varepsilon = c/2$ and $k > 1/\delta$. Using the value $\xi = x_k - x_0$ it follows from the Taylor expansion that

$$0 < c\,\|x_k - x_0\| \leq \|J_f(x_0)(x_k - x_0)\| = \|R_{x_0}(x_k - x_0)\| \leq \frac{c}{2}\,\|x_k - x_0\| \,.$$

Note that we used the fact that $f(x_k) = f(x_0) = p$. However, the above inequality is clearly a contradiction, so we conclude that $f^{-1}(p)$ is indeed a finite set. □

With this result we are ready for the first definition of the $\mathcal{C}^1$ mapping degree in the case of regular values.

Definition 17. Suppose that $f \in \mathcal{C}^1(\Omega, \mathbb{R}^n)$ where $\Omega \subseteq \mathbb{R}^n$ is open and bounded, and let $p \notin f(\partial\Omega)$ be a regular value. Then the $\mathcal{C}^1$ mapping degree is defined by

$$\deg(f, \Omega, p) = \sum_{x \in f^{-1}(p)} \operatorname{sign}(\det(J_f(x)))$$

The previous lemma ensures that the above sum is finite, and therefore that the construction is well-defined. As mentioned previously, the mapping degree can be seen as a sort of generalization or alternative to the the normal degree for polynomials. For instance, a statement such that $\deg(f, \Omega, p) = d$ immediately implies that $\#f^{-1}(p) \geq d$. We highlight a similarly obvious result, which will be useful in our later applications:

Corollary 1. *Let f be as above, and suppose that $\deg(f, \Omega, y) = \pm 1$. Then the equation $f(x) = y$ has an odd and positive number of solutions (i.e. the number $\#f^{-1}(y) \in \mathbb{N}$ is odd).*

However, there are some obstacles with this construction, the most obvious being that direct computation of the degree requires full knowledge on the number of solutions to the given equations, seemingly resulting in no new information. As will we see later, the way to overcome this issue is the so called homotopy invariance of the mapping degree, which in many practical cases allow us to compute the degree of a more straightforward function than initially given.

Another immediate issue is that naive applications of the mapping degree often yield less information than through elementary means. To illustrate this, consider the polynomial function

$$\begin{aligned} f : \mathbb{R} &\to \mathbb{R} \\ x &\mapsto x^2 - a, \end{aligned}$$

where $a > 0$. The degree of the polynomial is 2, so ideally we would want the mapping degree at 0 to also be 2 in the open ball $B_{\sqrt{a}+1}(0)$, which would account for the existence of the two roots $\pm\sqrt{a}$. However, we see that

$$\deg(B_{\sqrt{a}+1}(0), \mathbb{R}, 0) = f'(\sqrt{a}) + f'(-\sqrt{a}) = 0,$$

which provides no information about the number of roots. This issue will also partially be resolved later, as we will apply the mapping degree in a way that produces an alternative proof of the fundamental theorem of algebra. However this will be of mostly theoretical interest, since we are only interested in real positive roots for our later applications. Before moving on, we highlight a few useful properties of the mapping degree:

Normalization. Consider the identity $\mathrm{id} : \Omega \to \Omega$. We have that $J_{\mathrm{id}}(p) = I$ and that $\mathrm{id}^{-1}(p) = p$ for all $p \in \Omega$, so it clearly follows that

$$\deg(\mathrm{id}, \Omega, p) = 1.$$

Translation invariance. For any regular value p of f it clearly holds that 0 is a regular value of $f - p$, and furthermore that $J_{f-p}(x) = J_f(x)$. It therefore follows that

$$\deg(f - p, \Omega, 0) = \sum_{x \in (f-p)^{-1}(0)} \operatorname{sign}(\det(J_{f-p}(x))) = \sum_{x \in f^{-1}(p)} \operatorname{sign}(\det(J_f(x))) = \deg(f, \Omega, 0).$$

Additivity. Suppose that $\Omega_1, \Omega_2 \subseteq \Omega$ are disjoint subsets such that $p \notin f(\overline{\Omega} \backslash (\Omega_1 \cup \Omega_2))$. In particular this means that $f^{-1}(p) \subseteq \Omega_1 \cup \Omega_2$ so by definition we must have that

$$\deg(f, \Omega, p) = \deg(f, \Omega_1, p) + \deg(f, \Omega_2, p).$$

We now work towards showing the homotopy invariance of the mapping degree. First, we show that the mapping degree is invariant with respect to sufficiently small perturbations of p in the following sense:

Proposition 6. *Let $p \notin f(\mathrm{cl}(\Omega))$ be a regular value. There exists $\varepsilon > 0$ such that $B_\varepsilon(p)$ consists of only regular values and such that $\deg(f, \Omega, \cdot)$ is constant on $B_\varepsilon(p)$.*

Proof. This is an application of the inverse function theorem. For all $x \in f^{-1}(p)$ we find open neighborhoods $x \in W_x$, $p \in V_x$ such that the restriction $f|_{W_x} : W_x \longrightarrow V_x$ is a diffeomorphism. As $f^{-1}(p)$ is a finite set it holds that the intersection

$$\bigcap_{x \in f^{-1}(p)} V_x$$

is an open neighborhood of p, so it contains an open ball $B_\varepsilon(p)$ for some $\varepsilon > 0$. From the above diffeomorphisms we extract open subneigborhoods $x \in U_x \subseteq W_x$ that are diffeomorphic to $B_\varepsilon(p)$ via f, which means that all points $p' \in B_\varepsilon(p)$ are regular. Additionally,

$$f^{-1}(p') \subseteq \bigcup_{x \in f^{-1}(p)} U_x,$$

and each solution must be unique in U_x. Since $\det(J_f(\cdot))$ is continuous, its sign must be constant on U_x. This means that

$$\deg(f, \Omega, p') = \sum_{x' \in f^{-1}(p')} (\det J_f(x')) = \sum_{x \in f^{-1}(p)} (\det J_f(x)) = \deg(f, \Omega, p),$$

which was the desired result. □

Similarly, one can show that the mapping degree is invariant with respect to sufficiently small perturbations in f. To explain this terminology, let $\|\cdot\|_\infty$ denote the standard maximum norm, $f \in \mathcal{C}^1(\overline{\Omega}, \mathbb{R}^n)$, and recall that

$$\|f\|_{\mathcal{C}^1} := \|f\|_\infty + \max_{1 \le i \le n} \left\| \frac{\partial f}{\partial x_i} \right\|_\infty$$

defines a norm on the space $\mathcal{C}^1(\overline{\Omega}, \mathbb{R}^n)$.

Proposition 7. *Let $p \notin f(\partial\Omega)$ be a regular value. There exists $\varepsilon > 0$ such that for all functions $g \in \mathcal{C}^1(\overline{\Omega}, \mathbb{R}^n)$ where p is a regular value and $\|f - g\|_{\mathcal{C}^1} < \varepsilon$ it holds that $p \notin g(\partial\Omega)$, p is a regular value for g, and*

$$\deg(f, \Omega, p) = \deg(g, \Omega, p).$$

Proof. Our method for proving this is to establish a bijection between the sets $f^{-1}(p)$ and $g^{-1}(p)$ that preserves the sign of the respective Jacobian matrices. First, pick some $0 < \varepsilon < d(p, f(\Omega)$ smaller than the distance between p and the image $f(\Omega)$, and suppose that $g \in \mathcal{C}^1(\overline{\Omega}, \mathbb{R}^n)$ satisfies $\|f - g\|_{\mathcal{C}^1} < \varepsilon$. Then for all $x \in \partial\Omega$ it holds that

$$\|p - g(x)\| \ge \|\|p - f(x)\| - \|f(x) - g(x)\|\| \ge d(p, f(\partial\Omega)) - \varepsilon > 0,$$

which means that $p \notin g(\partial\Omega)$. Now, for all $y \in g^{-1}(p)$ we have that

$$p' := f(y) = p + (f(y) - g(y)) \in B_\varepsilon(p),$$

Similarly as in the proof of the previous proposition, we can for all $x \in f^{-1}(p)$ find open neighborhoods $x \in W_x$, $p \in V_x$ such that the restriction $f|_{W_x} : W_x \longrightarrow V_x$ is a diffeomorphism. By potentially letting $\varepsilon > 0$ be smaller, we can extract open subneighborhoods $x \in U_x \subseteq W_x$ that are diffeomorphic to $B_\varepsilon(p)$ via f. By assumption this means that

$$g^{-1}(p) \subseteq \bigcup_{x \in f^{-1}(p)} U_x.$$

Since p is a regular value for f there exists $c > 0$ such that $\|f'(x)\xi\| \ge c\,\|\xi\|$ for all $x \in f^{-1}(p)$ and $\xi \in \mathbb{R}^n$, and by continuity there exists $\delta > 0$ such that

$$\|J_f(y)\xi\| \ge \frac{c}{2}\,\|\xi\| \text{ for all } y \in \bigcup_{x \in f^{-1}(p)} B_\delta(x).$$

Now, we can pick $\varepsilon > 0$ such that $U_x \subseteq B_\delta(x)$ for all $x \in f^{-1}(p)$ and such that $\|f - g\|_{\mathcal{C}^1} < c/4$. Combining all this we get that

$$\|J_g(y)\xi\| \ge \|\|J_f(y)\xi\| - \|(J_f(y) - J_g(y))\xi\|\| \ge \frac{c}{4}\xi \text{ for all } y \in \bigcup_{x \in f^{-1}(p)} U_x,$$

which means exactly that p is a regular value for g. Note that the above inequality also gives a uniform bound $0 < 1/\|J_f(x)\| \le c_0$ for all $x \in f^{-1}(p)$. Let $x \in f^{-1}(p)$ and consider the remainder function $R_x(\xi) := g(x+\xi) - g(x) - J_g(x)\xi$. Define another function as follows

$$T(\xi) := [J_g(x)]^{-1}(p - g(x) - R_x(\xi)).$$

We wish to show that T has a unique fixed point in $B_\delta(0)$, which for $\xi = y - x$ implies the existence of a unique $y \in g^{-1}(p) \cap B_\delta(x)$. Observe that

$$\begin{aligned} R_x(\xi) - R_x(\xi') &= g(x+\xi) - g(x+\xi') - J_g(x)(\xi - \xi') \\ &= \int_{[0,1]} \big((J_g(x + t\xi + (1-t)\xi') - J_g(x)\big)(\xi - \xi')\,\mathrm{dt}. \end{aligned}$$

The triangle inequality gives a crude estimate of the integrand:

$$\begin{aligned} \|J_g(x + t\xi + (1-t)\xi') - J_g(x)\| &\le \|J_g(x + t\xi - (1-t)\xi') - J_f(x + t\xi + (1-t)\xi')\| \\ &\quad + \|J_f(x + t\xi + (1-t)\xi') - J_f(x)\| + \|J_f(x) - J_g(x)\| \\ &\le \|J_f(x + t\xi + (1-t)\xi')\| + 2\varepsilon. \end{aligned}$$

Combining this with the uniform bound $0 < 1/\|g'(x)\| \le c_0$, we get that

$$\begin{aligned} \|T(\xi) - T(\xi')\| &\le c_0 \|R_x(\xi) - R_x(\xi')\| \\ &\le c_0\big(\|J_f(x + t\xi + (1-t)\xi') - J_f(x)\| + 2\varepsilon\big) \|\xi - \xi'\|. \end{aligned}$$

Using the special case $\xi' = 0$ we derive the norm inequality

$$\begin{aligned} \|T(\xi)\| &\le c_0\big(\|J_f(x + t\xi) - f'(x)\| + 2\varepsilon\big)\|\xi\| + c_0\|p - g(x)\| \\ &\le c_0\big(\|J_f(x + t\xi) - f'(x)\| + 2\varepsilon\big)\|\xi\| + c_0\varepsilon, \end{aligned}$$

where we used that $\|p - g(x)\| = \|f(x) - g(x)\| < \varepsilon$. Now, from uniform continuity of J_f we get that there exists $\delta' > 0$ such that $\|J_f(x + t\xi + (1-t)\xi') - J_f(x)\| < 1/(4c_0)$ for all $t\|\xi\| + (1-t)\|\xi'\| < \delta'$. For $\delta'' < \min(\delta, \delta')$ and $\varepsilon < \min(1/(4c_0), \delta''/(4c_0))$ such that $U_x \subseteq B_{\delta''}(x)$ for all $x \in f^{-1}(p)$ we get that

$$\begin{aligned} &\|T(\xi) - T(\xi')\| \le \frac{3}{4}\|\xi - \xi'\|, \\ &\|T(\xi)\| \le \delta'' \text{ for all } \xi \in B_{\delta''}(0). \end{aligned}$$

This means that T is a contraction mapping, which we know has a unique fixed point (see exercise 28.7 in [7]). This establishes the desired bijection. The uniform bound $1/\|J_g(x)\| < c_0$ guarantees that the $\operatorname{sign}(J_g(\cdot))$ is constant in U_x, and we therefore conclude that

$$\deg(f, \Omega, p) = \sum_{x \in f^{-1}(p)} (\det J_f(x)) = \sum_{y \in g^{-1}(p)} (\det J_g(y)) = \deg(g, \Omega, p),$$

which was the desired result. □

We now work towards proving the homotopy invariance of the mapping degree. This is the most important nontrivial property of the degree, and requires some machinery. For starters, we recall the following standard results from analysis

Theorem 2. *[Approximation by smooth functions] Let $\Omega \subseteq \mathbb{R}^n$ be bounded and suppose $f \in \mathcal{C}^\ell(\overline{\Omega}, \mathbb{R}^m)$ and that f is $\mathcal{C}^{\ell'}$ on some open subset $\Omega' \subseteq \Omega$. Then there exists $(f_k)_{k\in\mathbb{N}} \subseteq \mathcal{C}^\infty(\overline{\Omega}, \mathbb{R}^m)$ such that*

$$\begin{aligned} &\|f_k - f\|_{\mathcal{C}^\ell} \to 0 \text{ on } \overline{\Omega}, \\ &\|f_k - f\|_{\mathcal{C}^{\ell'}} \to 0 \text{ on } \overline{\Omega'}. \end{aligned}$$

Theorem 3. *[Sard's theorem, a special case] Let $f \in \mathcal{C}^1(\Omega, \mathbb{R}^n)$, where $\Omega \subseteq \mathbb{R}^n$ is open. Then the set of regular values of f is dense in $\mathbb{R}^n$.*

For the sake of cohesion we omit the proof of both of these theorems; the interested reader may look towards appendix A in [10]. Having established these results we are ready to prove the first special case of homotopy invariance, namely by translating with a suitable constant.

Lemma 5. *Let $f \in \mathcal{C}^1(\overline{\Omega}, \mathbb{R}^n)$ and let p be a regular value such that $tp \notin f(\partial\Omega)$ for all $t \in [0,1]$. If $f^{-1}(0) = \emptyset$, then $\deg(f, \Omega, p) = \deg(f, \Omega, 0)$.*

Proof. By combining Theorem 2 with Proposition 7 we may freely assume that f is smooth. In particular, it follows from Theorem 3 that there exists a sequence $(\varepsilon_k)_{k\in\mathbb{N}} \subseteq \Omega$ of regular values for f and such that $\varepsilon_k \to 0$. Now, consider the auxiliary function

$$F : [-1,1] \times \overline{\Omega} \to \mathbb{R}^n$$
$$(s, x) \mapsto f(x) - (1 - s^2)p,$$

which is then also smooth by construction. Additionally, observe that $(0, \varepsilon_k)$ is a regular value for F for all $k \in \mathbb{N}$. Now, by the assumptions $tp \notin f(\partial\Omega)$ and $f^{-1}(0)$ we have that

$$0 \notin F(\partial([-1,1] \times \overline{\Omega}),$$

and that $F^{-1}(0) \subseteq (-1,1) \times \Omega$ is compact. By continuity of F this also means that the preimages $F^{-1}((0, \varepsilon_k))$ are also compact for all $k \geq N$ where $N \in \mathbb{N}$ is sufficiently large. In particular this means that these preimages can be written as a finite union of connected components. Since $(0, \varepsilon_k)$ is regular it follows from the implicit function theorem that $F^{-1}((0, \varepsilon_k))$ is the image of a smooth function. Therefore, the connected components must all be homeomorphic to embedded circles $\gamma \in (-1,1) \times \Omega$. Now, consider the hyperplane $\{s = 0\} \subseteq (-1,1) \times \Omega$ and note that

$$F^{-1}((0, \varepsilon_k)) \cap \{s = 0\} = f^{-1}(p + \varepsilon).$$

The symmetry $t \mapsto -t$ implies that every circle $\gamma^{-1}((0, \varepsilon_k))$ that intersects $\{s = 0\}$ must do so transversely, i.e. at exactly 2 points. This means that $f^{-1}(p + \varepsilon_k)$ must consist of an even (in particular finite) number of points. Write

$$\gamma \cap \{s = 0\} = \{(0, x^-), (0, x^+)\},$$

and note that γ is orientable. We therefore pick an oriented basis $\{-e_1, \xi^-\}$, $\{e_1, \xi^+\}$ at $(0, x^-)$, $(0, x^+)$ respectively. Applying the map J_F to this basis yields a basis of $\mathbb{R}^n$, but since J_F is continuous and $F|_\gamma = \varepsilon_k$ is constant, the orientation will be preserved, meaning that

$$J_F(0, x^{\pm})(\pm e) = 0, \qquad J_F(0, x^-)(\xi^-) = J_F(0, x^+)(\xi^+).$$

Note that $J_F(0, x^{\pm}) = (0\, J_f(x^{\pm})$, which means that $J_f(x^+) = -J_f(x^-)$, which immediately gives the desired result. □

We first prove the homotopy invariance in the special case where the homotopy is continuously differentiable.

Lemma 6. *Let $H : \overline{\Omega} \times [0,1] \to \mathbb{R}^n$ be a $\mathcal{C}^1$ homotopy such that $p \notin H(\partial\Omega \times [0,1])$. If p is a regular value for both $f_0 := H(\cdot, 0)$ and $f_1 := H(\cdot, 1)$, then*

$$\deg(f_0, \Omega, p) = \deg(f_1, \Omega, p).$$

Proof. Let $w : [-1,1] \to \mathbb{R}$ be a smooth function satisfying $w(s) = 0$ for all $s \in [-1, -1/2]$, $w(s) = 1$ for all $s \in [1/2, 1]$, and $w'(s) > 0$ for all $s \in (-1/2, 1/2)$. Consider the auxilliary function

$$G : [-1,1] \times \overline{\Omega} \to \mathbb{R}^n \times \mathbb{R} \cong \mathbb{R}^{n+1}$$
$$(s, x) \mapsto \begin{pmatrix} H(x, w(s)) - p \\ 4s^2 + 1 \end{pmatrix}.$$

By construction we have that $G \in \mathcal{C}^1([-1,1] \times \overline{\Omega})$. Additionally, for all $|s| \geq \frac{1}{2}$ we have that

$$\operatorname{sign}(\det J_G(s,x)) = \operatorname{sign}(J_{H(x,w(s))(x)} \cdot \operatorname{sign}(8s).$$

Now, consider $P := ((0, \dots, 0), 2) \in \mathbb{R}^{n+1}$. We have that

$$G^{-1}(P) = -\frac{1}{2} \times f_0^{-1}(p) \cup \frac{1}{2} \times f_1^{-1}(p),$$

which in particular means that $|s| \geq \frac{1}{2}$ for all $(s,x) \in G^{-1}(P)$, so P is a regular value for G. By construction we have that

$$\begin{aligned}\deg(G,(1,1)\times\Omega,P) &= \deg(f_1-p,\Omega,0)-\deg(f_0-p,\Omega,0)\\ &= \deg(f_1,\Omega,p)-\deg(f_0,\omega,p).\end{aligned}$$

Now, for $t \in [0,1/2)$ we have that $G^{-1}(tP)=\emptyset$ and for $t\in[1/2,1]$ we have that

$$G^{-1}(tP) = -\frac{\sqrt{2t-1}}{2}\times f_0^{-1}(p)\cup\frac{\sqrt{2t-1}}{2}\times f_1^{-1}(p)\subseteq(-1,1)\Omega$$

In particular $G^{-1}(0,0)=\emptyset$, so from Lemma 5 we conclude that

$$\deg(G,(1,1)\times G,P)=0,$$

which proves the desired result. $\square$

With these two lemmas in place we are finally able to formulate and prove the general homotopy invariance of the mapping degree:

Proposition 8. *Let $H:\overline{\Omega}\times[0,1]\to\mathbb{R}^n$ be a homotopy such that $p\notin H(\partial\Omega\times[0,1])$. If p is a regular value for both $f_0 := H(\cdot,0)$ and $f_1 = H(\cdot,1)$ then*

$$\deg(f_0,\Omega,p)=\deg(f_1,\Omega,p).$$

Proof. By using a suitable reparametrization we can freely assume that $H(x,t)=H(x,0)=f_0(x)$ for all $t\in[0,1/3]$ and that $H(x,t)=H(x,1)=f_1(x)$ for all $t\in[2/3,1]$. Note that this in particular means that H is $\mathcal{C}^1$ on the open set $(0,1/3)\cup(2/3,1)$. From Theorem 2 we know that there exists a sequence $\widetilde{H_k}\in\mathcal{C}^\infty(\overline{\Omega}\times[0,1])$ such that:

$$\begin{aligned}\left\|H(x,t)-\widetilde{H_k}(x,t)\right\|_{\mathcal{C}^0}&\to 0\quad\forall(x,t)\in\overline{\Omega}\times[1/3,2/3],\\ \left\|H(x,t)-\widetilde{H_k}(x,t)\right\|_{\mathcal{C}^1}&\to 0\quad\forall(x,t)\in\overline{\Omega}\times((0,1/3)\cup(2/3,1)).\end{aligned}$$

Writing $\widetilde{f_{0,k}}(x):=\widetilde{H_k}(x,0)$ and $\widetilde{f_{1,k}}(x):=\widetilde{H_k}(x,1)$, then the second limit above means in particular that $\widetilde{f_{0,k}}\to f_0$ and $\widetilde{f_{1,k}}\to f_1$ in $\mathcal{C}^1$. Choosing $k\geq N_1$ for some $N_1\in\mathbb{N}$ large enough such that $p\notin H_k(\Omega\times[0,1])$ means that the conditions necessary to invoke Lemma 6 are satisfied, and we therefore have that

$$\deg(\widetilde{f_{0,k}},\Omega,p)=\deg(\widetilde{f_{1,k}},\Omega,p).$$

Choosing $k\geq N_2\geq N_1$ for some $N_2\in\mathbb{N}$ such that both $\widetilde{f_{0,k}}$ and $\widetilde{f_{1,k}}$ are small enough pertubations in $\mathcal{C}^1$ of f_0 and f_1 respectively, it follows from the pertubation invariance of the mapping degree that

$$\deg(f_0,\Omega,p)=\deg(\widetilde{f_{0,k}},\Omega,p)=\deg(\widetilde{f_{1,k}},\Omega,p)=\deg(f_1,\Omega,p),$$

which was the desired result. $\square$

5.2 A Few Other Applications

We include this section simply to concretely demonstrate that the $\mathcal{C}^1$-mapping degree allows us to prove useful results about the number of solutions for equations of a specific type.

Proposition 9. *Suppose that $D \subseteq \mathbb{R}^n \setminus f(\partial\Omega)$ is a connected component (for instance D is convex), and consider a path $p : [0,1] \to D$ such that both $p(0)$ and $p(1)$ are regular values of f. Then it holds that*

$$\deg(f, \Omega, p(0)) = \deg(f, \Omega, p(1)).$$

Proof. The path gives rise to a homotopy between $f - p(0)$ and $f - p(1)$ as follows:

$$\begin{aligned} H : \overline{\Omega} \times [0,1] &\to \mathbb{R}^n \\ (x,t) &\mapsto f(x) - p(t). \end{aligned}$$

Note that since $p(t) \notin f(\partial\Omega)$ for all $t \in [0,1]$ it follows that $0 \notin H(\partial\Omega \times [0,1])$, so from the homotopy invariance we get that

$$\deg(f - p(0), \Omega, 0) = \deg(f - p(1), \Omega, 0),$$

which was the desired result. $\square$

Corollary 2 (Fundamental Theorem of Algebra)**.** *Let $f : \mathbb{C} \to \mathbb{C}$ be a polynomial of degree d. Considering f as a map $f : \mathbb{R}^2 \to \mathbb{R}^2$ we have that $\deg(f, B_R(0), 0) = d$ for some $R > 0$.*

Proof. Note that $\deg(f, B_R(0), 0) = d$ implies that there are at least d roots in the open set $B_r(0)$, so by factorization of the polynomial we conclude that these must be the total number of roots, which is exactly the fundamental theorem of algebra. We can write

$$f(z) = az^d + q(z)$$

where $a \neq 0$ and q is some polynomial of degree less than d. Now, consider the following homotopy

$$\begin{aligned} H : \mathbb{R}^2 \times [0,1] &\to \mathbb{R}^2 \\ (z,t) &\mapsto az^d + tq(z). \end{aligned}$$

Clearly H is continuous and is a homotopy between f and the function az^d. For all $z \in \partial B_R(0)$ where $R > 0$ is sufficiently large, there exists a $C > 0$ such that $|q(z)| \leq CR^{d-1}$. This means that and therefore also that

$$|H(z,t)| \geq |a|R^d - CR^{d-1} > 0,$$

for $R > 0$ sufficiently large. I

which means that $0 \notin H(\partial B_R(0) \times [0,1])$, so by the homotopy invariance of the degree it follows that

$$\deg(f, B_R(0), 0) = \deg(az^n, B_r(0), 0).$$

Now, we have that $B_R(0)$ is connected and that both 0 and a are regular values of az^n. By picking $R > 0$ large enough such that $a \in B_R(0)$ it follows from Proposition 9 that

$$\deg(az^n, B_r(0), 0) = \deg(az^n, B_r(0), a).$$

The equation $az^d = a$ has exactly d distinct solutions $z_1, \ldots, z_d$, and writing $az^d = f_1(z) + if_2(z)$ we see that

$$\det J_f(z_i) = \left(\frac{\partial f_1(z_i)}{\partial x_2}\right)^2 + \left(\frac{\partial f_2(z_i)}{\partial x_1}\right)^2 > 0,$$

where we used that f_1, f_2 satisfy the Cauchy-Riemann equations since az^d is holomorphic. We therefore conclude that

$$\deg(f, B_r(0), 0) = \deg(az^d, B_r(0), a) = d,$$

which was the desired result. $\square$

5.3 The $\mathcal{C}^1$-Mapping Degree for Nonregular Values

We now also briefly extend the $\mathcal{C}^1$-mapping degree to nonregular values.

Definition 18. Let $D \subseteq \mathbb{R}^n \setminus f(\Omega)$ be connected and $p \in D$. Then the $\mathcal{C}^1$-mapping degree is defined by

$$\deg(f, \Omega, p) := \lim_{k\to\infty} \deg(f, \Omega, p_k),$$

where $(p_k)_{k\in\mathbb{N}} \subseteq \mathbb{R}^n \setminus f(\partial\Omega)$ consists of regular values with $p_k \to p$.

Note that such a sequence $(p_k)_{k\in\mathbb{N}}$ exists by Sard's theorem. It follows from the perturbation invariance of the mapping degree for regular values that the construction is well-defined, and from Proposition 9 we see that the degree is invariant in D. Importantly, this slightly more general definition satisfies the same key properties that we desire.

Theorem 4. *The degree from Definition 18 satisfies the following properties:*

- **Normalization.** *If* $p \in \Omega$ *then* $\deg(\mathrm{id}, \Omega, p) = 1$.
- **Translation invariance.** $\deg(f, \Omega, p) = \deg(f - p, \Omega, 0)$.
- **Additivity.** *For disjoint open subsets* $\Omega_1, \Omega_2 \subseteq \Omega$ *with* $p \notin f(\overline{\Omega} \setminus (\Omega_1 \cup \Omega_2))$ *it holds that*

$$\deg(f, \Omega, p) = \deg(f, \Omega_1, p) + \deg(f, \Omega_2, p).$$

- **Homotopy invariance.** *Let* $H : \overline{\Omega} \times [0,1] : \mathbb{R}^n$ *be a homotopy between* f_0 *and* f_1 *and let* $p \notin H(\partial\Omega \times [0,1])$. *Then*

$$\deg(f_0, \Omega, p) = \deg(f_1, \Omega, p)$$

Proof. The proofs of each of the properties is more of less derived from the same properties of the mapping degree for regular values. Normalization in particular follows directly, since $p \in \Omega$ is a regular value for the identity.

To see additivity, suppose that $(p_k)_{k\in\mathbb{N}}$ is a sequence of regular values with $p_k \to p$. The condition that $f^{-1}(p) \subseteq \Omega_1 \cup \Omega_2$ implies in particular that $f^{-1}(p_k) \in \Omega_1 \cup \Omega_2$ for all $k \geq N$ with $N \in \mathbb{N}$ large enough. Furthermore, from pertubation invariance of the mapping degree for regular values we can assume that $\deg(f, \Omega, p_k)$ is unchanged for all $k \geq N$. So, from the additivity of the mapping degree for regular values we have that

$$\begin{aligned}
\deg(f, \Omega, p) &= \lim_{\substack{k\to\infty \\ k\geq N}} \deg(f, \Omega, p_k) \\
&= \lim_{\substack{k\to\infty \\ k\geq N}} (\deg(f\Omega_1, p_k) + \deg(f, \Omega_2, p_k)) \\
&= \deg(f, \Omega_1, p) + \deg(f, \Omega_2, p).
\end{aligned}$$

To see homotopy invariance, let $H : \overline{\Omega} \times [0,1]$ be homotopy with the assumed properties. We can pick a sequence $(p_k)_{k\in\mathbb{N}}$ of regular values for both f_0 and f_1 with $p_k \to p$. Now, there exists $\delta > 0$ such that $d(p, H(\partial\Omega \times [0,1])) \geq \delta$. This means that there exists $N \in \mathbb{N}$ such that $d(p_k, H(\partial\Omega \times [0,1])) \geq \delta/2$ for all $k \geq N$. In particular we then have that $p_k \notin H(\partial\Omega \times [0,1])$ for all $\geq N$. It then follows from the homotopy invariance of the mapping degree for regular values (Proposition 8) that

$$\deg(f_0, \Omega, p) = \lim_{\substack{k\to\infty \\ k\geq N}} \deg(f_0, \Omega, p_k) = \lim_{\substack{k\to\infty \\ k\geq N}} \deg(f_1, \Omega, p_k) = \deg(f_0, \Omega, p)$$

Lastly, to see translation invariance, let $p_k \to p$ consist of regular values and note that by translation invariance for regular values we have that

$$\deg(f, \Omega, p) = \lim_{k\to\infty} \deg(f, \Omega, p_k) = \lim_{k\to\infty} \deg(f - p_k, \Omega, 0)$$

Now, it is easily verified that the straight-line homotopy H_k between $f - p_k$ and $f - p$ satisfy that $0 \notin H_k(\partial\Omega \times [0,1])$ for all $k \geq N$ where $N \in \mathbb{N}$ is sufficiently large. We therefore conclude that

$$\deg(f, \Omega, p) = \lim_{\substack{k\to\infty \\ k\geq N}} \deg(f - p_k, \Omega, 0) = \deg(f - p, \Omega, 0),$$

which was the desired result. $\square$

In fact, it is possible to show that these four properties uniquely characterize the degree from Definition 18. In particular the $\mathcal{C}^1$-mapping degree is also unique. While this uniqueness property is usually not relevant when computing the degree for some specific function, it is still good to remember.

5.4 Degree for Functions of the Form φ_c

Recall that we defined the function $\varphi_c \in \mathcal{C}^1(\mathbb{R}^n_{\geq 0}, \mathbb{R}^n)$ coordinate-wise as

$$\varphi_c(x)_i = \begin{cases} f_i(x), & i \notin \{i_1, \dots, i_d\}, \\ (Wx - c)_i & i \in \{i_1, \dots, i_d\}, \end{cases}$$

where $f \in \mathcal{C}^1(\mathbb{R}^n_{\geq 0}, \mathbb{R}^n)$, $W \in \mathbb{R}^{d \times n}$ is a row-reduced matrix, and $c \in \mathbb{R}^d$. The first result about this function is that for open subsets satisfying certain regularity conditions, we have a nice expression for the degree. We first recall some basic notions from the theory of metric spaces convex sets.

Definition 19. Let $B \subseteq \mathbb{R}^n$ be a convex set and $v \in \mathbb{R}^n$ a point. We say that v points inwards B at $x \in \partial B$ if there exists $\varepsilon > 0$ such that $x + tv \in \overline{B}$ for all $t \in (0, \varepsilon)$. Conversely, we say that v points outwards B at $x \in \partial B$.

Example 5. A trivial instance of this occurrence is that $v = 0$ points inwards B on the whole boundary ∂B.

Example 6. Another important example is that for all $x_1 \in B$ and $x_2 \in \partial B$ we have that $x_1 - x_2$ points inwards B at x_2. If B is open we have a converse example, namely that $x_2 - x_1$ points outwards B at x_2. Both of these examples follow almost directly from the definition of convexity.

With this terminology in place we are ready to prove the main theorem regarding the degree of general functions of the form φ_c.

Theorem 5. *Let $f \in \mathcal{C}^1(\mathbb{R}^n_{\geq 0}, \mathbb{R}^n)$, $W \in \mathbb{R}^{d \times n}$ and φ_c be constructed as above. Suppose that $B_c \subseteq \mathbb{R}^n_{\geq 0}$ is open, bounded, convex, and satisfies*

1. *(i) $B_c \cap \mathcal{P}_c \neq \emptyset$.*
2. *(ii) $f(x) \neq 0$ and $Wf(x) = 0$ for all $x \in \partial B_c \cap \mathcal{P}_c$.*
3. *(iii) For all $x \in \partial B_c \cap \mathcal{P}_c$ the vector $f(x)$ points inwards B_c at x.*

Then it holds that $\deg(\varphi_c, B_c, 0) = (-1)^{n-d}$.

Proof. The method of the proof is to compute the mapping degree for a simpler function that is homotopic to φ_c. For this purpose, let $\bar{x} \in B_c \cap \mathcal{P}_c$ be fixed and define $G : \mathrm{cl}(B_c) \to \mathbb{R}^n$ by

$$G(x) = (Wx - c, \pi(\bar{x} - x)) \in \mathbb{R}^d \times \mathbb{R}^{n-d} \cong \mathbb{R}^n,$$

where $\pi : \mathbb{R}^n \to \mathbb{R}^s$ is the projection to the last $s := n - d$ coordinates. Recall from the discussion after Definition 10 that we can, without loss of generality, reorder the coordinates in a suitable way in order to get an equality $W = \begin{pmatrix} I_d & \widehat{W} \end{pmatrix}$ of block matrix. We then have that

$$J_G(x) = \begin{pmatrix} I_d & \widehat{W} \\ 0 & -I_s \end{pmatrix}.$$

It follows from linear algebra that $\det(J_G(x)) = \det I_d \det(-I_s) = (-1)^s$, so in particular this means that 0 is a regular value for G. It also follows from its definition that $G(x) = 0$ if and only if $Wx = W\bar{x} = c$ and $\pi(x) = \pi(\bar{x})$. However, using the block form of W once again, we conclude that this means that $x = \bar{x}$. In particular this means that G is non-zero on $\overline{B_c}$. It then follows from the definition of the $\mathcal{C}^1$ mapping degree that

$$\deg(G, B_c, 0) = \mathrm{sign}(\det(J_G(\bar{x}))) = (-1)^s.$$

Now, consider the straight-line homotopy between φ_c and G

$$\begin{aligned} H : \overline{B_c} \times [0,1] &\to \mathbb{R}^n \\ (x,t) &\mapsto t\varphi_c(x) + (1-t)G(x). \end{aligned}$$

We wish to show that H satisfy the criteria of Proposition 8, i.e. that $0 \notin H(\partial B_c \times [0,1])$. So, suppose for the sake of contradiction that $H(x', t) = 0$ for some $(x', t) \in \partial B_c \times [0,1]$. From the definitions of both φ_c and G we have that

$$H(x', t) = (Wx' - c, t\pi(f(x')) + (1-t)\pi(\bar{x} - x')).$$

So in particular we have $x' \in \mathrm{bd}(B_c) \cap \mathcal{P}_c$. The edge case $t = 0$ is equivalent to $G(x') = 0$ which we know is impossible. The case $t = 1$ means that $\pi(f(x')) = 0$, but since we have $Wf(x) = 0$ by assumption, this must mean that $f(x') = 0$ which is a contradiction. The only remaining possibility is then that $t \in (0,1)$. In this case we have that

$$\pi(f(x')) = \frac{t-1}{t}\pi(\bar{x} - x').$$

However, since $Wf(x') = W(\bar{x} - x') = 0$ it once again follows that

$$f(x') = \frac{t-1}{t}(\bar{x} - x').$$

From Example 6 we know that the difference $x' - \bar{x}$ points outwards B_c at x'. Since $(t-1)/t < 0$ we therefore conclude from the above equality that $f(x')$ points outwards B_c at x', which yet again contradicts our initial assumptions. We therefore conclude that $H(x,t) \neq 0$ for all $(x,t) \in \mathrm{bd}(B_c) \cap \mathcal{P}_c$, so H satisfy the criteria from Proposition 8, which means that

$$\deg(\varphi_c, B_c, 0) = \deg(G, B_c, 0) = (-1)^{n-d}.$$

□

A crucial property of the set B_c in the above theorem is that the function $f \in \mathcal{C}^1(\mathbb{R}^n_{\geq 0}, \mathbb{R}^n)$ is nonzero on the relevant parts of the boundary B_c. When we revisit this theorem later, then in the context of reaction networks, this property will translate into the formation rate function not having any boundary equilibria in some nontrivial stoichiometric compatibility class. It is therefore worth mentioning already that the theory about siphons will be very practical when applying this theorem to concrete networks.

It is also worth mentioning that the above proof used the homotopy invariance of the $\mathcal{C}^1$-mapping degree in quite an interesting way. In our proof of the fundamental theorem of algebra we constructed a suitable homotopy between an arbitrary polynomial $f(z)$ of degree n and the function az^n. This was a perfect situation, because we can directly compute all the roots of this function and therefore also the degree. In the above proof we had a function φ_c, which we showed was homotopic to a function G which was harder to concretely write down. However, here we had the theoretical assumptions necessary to compute the degree without knowing the concrete solutions to the relevant equation.

6 MAIN THEOREM: Multistationarity for Dissipative Networks

6.1 Degree in the Setting of Reaction Networks

Let G be a CRN with associated mass-action system $\dot{x} = f(x)$, and recall the convention $\dim_{\mathbb{R}}(S) = s$. Let $W \subseteq \mathbb{R}^{d\times n}$ be a row-reduced matrix whose rows form a basis of $S^{\perp}$. In this setting, we can construct the function $\varphi_c \colon \mathbb{R}^n_{\geq 0} \to \mathbb{R}^n$ in the same way as in the previous section. The goal is now to show the existence of the open, bounded and convex set B_c in this setting, which will require several assumptions that we are able to verify using methods developed earlier in the project.

Theorem 6. *Let G be a dissipative CRN with associated mass-action system $\dot{x} = f(x)$, and suppose that $\dim_{\mathbb{R}}(S) = s$ and that $W \in \mathbb{R}^{d\times n}$ is a row-reduced matrix whose rows form a basis of $S^{\perp}$. Additionally, suppose that $c \in \mathbb{R}^d$ is chosen such that $\mathcal{P}_c^+ \neq \emptyset$ and that $f(x) \neq 0$ for all $x \in \mathcal{P}_c^c ap \partial\mathbb{R}^n_{\geq 0}$. Then there exists an open, bounded, and convex set $B_c \subseteq \mathbb{R}^n_{\geq 0}$ such that $V(f) \cap \mathcal{P}_c^+ \subseteq B_c$ and such that*

$$\deg(\varphi_c, B_c, 0) = (-1)^s$$

Proof. The goal is to apply Theorem 5 in a suitable setting. Let $K_c \subseteq \mathcal{P}_c$ be a compact attracting set of all trajectories with initial condition in $\mathcal{P}_c$, and recall from Lemma 1 that we can freely assume that K_c is forward invariant, not contained in $\partial\mathbb{R}^n_{\geq 0}$, and contains all ω-limit points of $\mathcal{P}_c^+$ in its interior. Now, let $B \subseteq \mathbb{R}^n$ be an open, bounded, and convex set with $K_c \subseteq B$, and define $B_c = B \cap \mathbb{R}^n_{>0}$. We have that

$$\emptyset \neq K_c \cap \mathbb{R}^n_{>0} \subseteq B_c \cap \mathcal{P}_c.$$

Similarly, by using that $0 \notin f(\partial\mathbb{R}^n_{\geq 0})$, we see that

$$V(f) \cap \mathcal{P}_c^+ \subseteq K_c \cap \mathbb{R}^n_{>0} \subseteq B_c$$

We now work towards defining a function, which, together with the set B_c above, will satisfy the criteria in Theorem 5. For this purpose, let $U_1 \subsetneq B$ be an open set containing K_c and write $U_2 = \mathbb{R}^n \setminus K_c$. By construction, we have an open cover $\mathbb{R}^n = U_1 \cup U_2$ with $U_1 \cap U_2 \neq \emptyset$, so a simple topological argument (see for instance Theorem 36.3 in [7]) allows us to select a $\mathcal{C}^1$-partition of unity $\psi_1, \psi_2 \in \mathcal{C}^1(\mathbb{R}^n, [0,1])$. Let $\widetilde{x} \in K_c \cap \mathbb{R}^n_{>0}$ be arbitrary and fixed, and consider the auxiliary function

$$\begin{aligned} \rho : \mathbb{R}^n_{\geq 0} &\to \mathbb{R}^n \\ x &\mapsto \psi(x)(\widetilde{x} - x). \end{aligned}$$

We once again consider the function

$$\begin{aligned} \tau : \mathbb{R}^n_{\geq 0} &\to \mathbb{R}_{\geq 0} \\ x &\mapsto \inf\{t \geq 0 : \Phi(x,t) \in K_c\}, \end{aligned}$$

as well as the associated maximum entry time to K_c for trajectories starting in $\overline{B_c} \cap \mathcal{P}_c$:

$$T = \max\{\tau(x) \mid x \in \overline{B_c} \cap \mathcal{P}_c\}.$$

Recall that $\overline{B_c} \cap \mathcal{P}_c$ is compact, so from the exact same argument as in the proof of Lemma 1 it follows that $T < \infty$. In the case that $T = 0$, i.e. that $\tau(x) = 0$ for all $x \in \overline{B_c} \cap \mathcal{P}_c$ (for instance, if the network is conservative) we redefine $T = 1$. Now, consider the function

$$\begin{aligned} g : \mathbb{R}^n_{\geq 0} &\to \mathbb{R}^n \\ x &\mapsto \frac{1}{T}(\Phi(x,T) - x) + T\rho(x). \end{aligned}$$

From the definition of T it holds in particular that $\Phi(x,T) \in \overline{B_c} \cap \mathcal{P}_c$ for all $x \in \partial B_c \cap \mathcal{P}_c$, which means that

$$Wg(x) = W\left(\frac{1}{T}(\Phi(x,T) - x)\right) + W(T\rho(x)) = 0.$$

Furthermore, note that both $T\rho(x)$ and $\frac{1}{T}(\Phi(T,x)-x)$ point inwards B_c at $x \in \partial B_c \cap \mathcal{P}_c$ (where we in particular used that $T\psi_1(x) \geq 0$ in the definition of $\rho(x)$). Therefore, also the sum $g(x)$ points inwards B_c at $x \in \partial B_c \cap \mathcal{P}_c$. From Theorem 5 it then follows that the function φ_c^g constructed from g and W satisfies

$$\deg(\varphi_c^g, B_c, 0) = (-1)^s.$$

Now, consider the unnormalized homotopy

$$H(x,t) = \begin{cases} \varphi_c(x), & t = 0, \\ (Wx - c, \frac{1}{t}\pi(\Phi(x,t) - x) + t\pi(\rho(x)), & t \in (0,T]. \end{cases}$$

Note that $\Phi(x,t)$ is differentiable, which in particular means that H is continuous. Additionally, $H(x,T) = (Wx - c, \pi(g(x))) = \varphi_c^g$, which means that H is indeed a homotopy between φ_c and φ_c^g. We therefore need to show that $H(x,t) \neq 0$ for all $(x,t) \in \partial B_c \cap [0,T]$ in order to use the homotopy invariance of the mapping degree. First note that we cannot have $H(x,0) = \varphi_c(x) = 0$ for any $x \in \partial B_c$ since $V(f) \cap \mathcal{P}_c \subseteq B_c$. So, let $x' \in \partial B_c$ and suppose for the sake of contradiction that $H(x',t)$ for some $t \in (0,T]$. By definition of H we get that $x' \in \partial B_c \cap \mathcal{P}_c$ and that

$$\pi(\Phi(x',t) - x') = -t^2\pi(\rho(x)).$$

However, since $W\rho(x') = 0$ we know that $\Phi(x',t) = x' - t^2\rho(x')$. Note that $\rho(x') = 0$ then by definition we must have $\psi_1(x') = 0$. Additionally, by the above equation x' must be a fixed point and thus contained in K_c. However, this means that $\psi_1(x') = 1 \neq 0$, which is clearly a contradiction, so we must have $\rho(x') \neq 0$, and therefore in particular $\psi_1(x') \neq 0$, which means that $x' \in U_1$. Note that

$$\mathbb{R}^n_{>0} \cap \partial B_c \cap U_1 \subseteq \partial B \cap U_1 = \emptyset,$$

which means that $x' \in \partial R^n_{\geq 0}$, i.e. that $x_i = 0$ for some $i \leq n$. From the previous equation this means that

$$\Phi(x',t)_i = x'_i - t^2\rho(x')_i = -t^2\psi_1(\widetilde{x} - x')_i = -t^2\psi_1(x')\widetilde{x}_i < 0.$$

However, this contradicts the forward invariance of the positive orthant $\mathbb{R}^n_{\geq 0}$, so we conclude that $H(x',t) \neq 0$. Using the homotopy invariance of the mapping degree we therefore conclude that

$$\deg(\varphi_c, B_c, 0) = \deg(\varphi_c^g, B_c, 0) = (-1)^s,$$

which was the desired result. □

The partial derivatives of φ_c are independent of the value of $c \in \mathbb{R}^d$, which allows for the definition $M(x) = J_{\varphi_c}(x)$. We have that

$$M(x)_i = \begin{cases} J_{f_i}(x), & i \notin \{i_1, \ldots, i_d\} \\ W_i, & i \in \{i_1, \ldots, i_d\} \end{cases}$$

Definition 20. A steady state $x' \in \varphi_c^{-1}(0)$ is said to be non-degenerate if $M(x)'$ is invertible, i.e. if $\det M(x') \neq 0$.

Now that we have shown the existence of the crucial set B_c, we can finally state the following theorem, which is in a way the culmination of our applications of the $\mathcal{C}^1$-mapping degree in the setting of reaction networks.

Theorem 7. *Suppose that f is the formation rate function for some mass-action network with stoichiometric subspace S and let $\mathcal{P}_c$ be a stoichiometric compatibility class such that $\mathcal{P}_c^+ \neq 0$. If the network is dissipative and $f(x) \neq 0$ for all $x \in \mathrm{bd}(\mathbb{R}^n_{\geq 0}) \cap \mathcal{P}_c$, then we have the following criteria:*

- *If sign$((-1)^s \det M(x)) = 1$ for all $x \in V(f) \cap \mathcal{P}_c^+$, then $\#V(f) \cap \mathcal{P}_c^+ = 1$, with this unique steady state also being non-degenerate.*
- *If sign$((-1)^s \det M(x)) = -1$ for some $x \in V(f) \cap \mathcal{P}_c^+$ then $\#V(f) \cap \mathcal{P}_c^+ \geq 2$ and at least one of these steady states will be non-degenerate.*

Proof. It follows from Theorem 6 that we can pick an open bounded convex set $B_c \subseteq \mathbb{R}^n_{>0}$ such that

$$V(f) \cap \mathcal{P}_c \subseteq B_c, \qquad \deg(\varphi_c, B_c, 0) = (-1)^s.$$

Note that we have an equality $V(f) \cap \mathcal{P}_c^+ = B_c \cap V(\varphi_c)$. Now, for the first part of the theorem assume that $\text{sign}(\det M(x)) = (-1)^s \neq 0$ for all $x \in V(f) \cap \mathcal{P}_c$. In particular, this means that 0 is a regular value for φ_c, so taking the mapping degree yields an equality

$$(-1)^s = \deg(\varphi_c, B_c, 0) = \sum_{x \in B_c \cap V(\varphi_c)} \text{sign}(\det M(x)) = (-1)^s \# B_c \cap V(\varphi_c).$$

From this we easily deduce that $\# B_c \cap V(\varphi_c) = 1$, so there is a unique steady state (which clearly is non-degenerate by assumption) in the stoichiometric compatibility class P_c.

For the second part of the theorem, assume that $\text{sign}(\det M(y)) = (-1)^{s+1}$ for some $y \in V(f) \cap \mathcal{P}_c^+$. In the case where 0 is a regular value for φ_c, then similarly to above it follows from the mapping degree that we have an equality

$$(-1)^s = \deg(\varphi_c, B_c, 0) = \sum_{x \in B_c \cap V(\varphi_c)} \text{sign}(\det M(x)) = (-1)^{s+1} + \sum_{\substack{x \in B_c \cap V(\varphi_c) \\ x \neq y}} \text{sign}(\det M(x)).$$

From this we conclude that there must exist at least two other points $x', x'' \in B_c \cap V(\varphi_c)$ such that

$$\text{sign}(\det M(x')) = \text{sign}(\det M(x'')) = (-1)^s.$$

Note that in particular these points are also non-degenerate. Now, suppose that 0 is not a regular value for φ_c. By definition this means that there exists $x' \in B_c \cap V(\varphi_c)$ such that $\det M(x') = 0$. In other words, $\mathcal{P}_c$ has at least two positive steady states, and by assumption one of these is non-degenerate. $\square$

Let $m < n$ and suppose that we have a function $\Phi : \mathbb{R}^m \to \mathbb{R}^n$ satisfying $\Phi(\mathbb{R}^m_{>0}) \cong V_{>0}(f)$ (in other words, $V_{>0}(f)$ is positively parameterized by Φ). In this case, we have a bijection

$$V_{>0}(f) \cap \mathcal{P}_c \longleftrightarrow \mathcal{S}_c := \{\hat{x} \in \mathbb{R}^m_{>0} \mid W\Phi(\hat{x}) = c\}$$

We define $a(\hat{x}) := \det(M(\Phi(x)))$, which allows for a slight generalization of Theorem 7, which also happens to be more practical.

Corollary 3. *In the same setting as in Theorem 7, we have the following criteria:*

- *If sign*$((-1)^s a(\hat{x})) = 1$ *for all* $\hat{x} \in \mathbb{R}^m_{>0}$*, then* $|V(f) \cap \mathcal{P}_c^+| = 1$ *for all* $\mathcal{P}_c$ *such that* $\mathcal{P}_c^+ \neq \emptyset$*. Furthermore, this unique equilibria is non-degenerate.*
- *If sign*$((-1)^s a(\hat{x})) = -1$ *for some* $\hat{x} \in \mathbb{R}^m_{>0}$ *then* $|V(f) \cap \mathcal{P}^+_{W\Phi(\hat{x})}| \geq 2$*, with at least one of these being non-degenerate. If all equilibria in* $\mathcal{P}_{W\Phi(\hat{x})}$ *are non-degenerate, then the number of equilibria is at least three and always odd.*

Proof. Suppose that $c > 0$ is chosen such that $\mathcal{P}_c^+$. For all $x \in V_{>0}(f) \cap \mathcal{P}_c$ there exists a unique $\hat{x} \in \mathbb{R}^m$ such that $x = \Phi(\hat{x})$, and we have that

$$\det(M(x)) = \det(M(\Phi(\hat{x}))) = a(\hat{x})$$

The corollary then clearly follows from Theorem 7. $\square$

6.2 Application to Specific Reaction Networks

We now apply the above results to a number of concrete reaction networks, some of which have been previously discussed. In order to apply Corollary 3 we will frequently use the following fundamental property of the Newton Polytope $N(f)$, which is defined as the convex hull of the exponent vectors for a given polynomial f.

Proposition 10. *Let $f(x) = \sum_{\alpha \in \mathbb{N}^n} c_\alpha x^\alpha \in \mathbb{R}[x_1, \dots, x_n]$ and let α' be a vertex of $N(f)$. Then there exists $x' \in \mathbb{R}^n_{>0}$ such that $\mathrm{sign}(f(x')) = \mathrm{sign}(c_{\alpha'})$.*

The proof follows from basic theory on polytopes and is omitted.

Example 1, continued. Recall that we have shown this network to be both conservative and devoid of any boundary steady states in any S-class $\mathcal{P}_c$ with $\mathcal{P}_c^+ \neq \emptyset$. Now, let $c \in \mathbb{R}^2$ be arbitrary. We have that the function φ_c constructed from f and W is given by

$$\varphi_c(x_1, x_2, x_3, x_4) = \begin{pmatrix} x_1 - c_1 \\ x_2 - c_2 \\ k_1 x_1 - k_3 x_3 \\ k_2 x_2 - k_4 x_4 \end{pmatrix},$$

and the Jacobian matrix $M(x) = J_{\varphi_c}(x)$ is

$$M(x) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ k_1 & 0 & -k_3 & 0 \\ 0 & k_2 & 0 & -k_4 \end{pmatrix}.$$

A simple computation shows that the determinant is equal to $k_3 k_4 > 0$, so from Theorem 7 we conclude that there is a unique, non-degenerate steady state for the system. This should also be obvious from the fact that φ_c consists only of linear functions, and using the conservation laws of the network we see that this solution is explicitely given by

$$(x_1, x_2, x_3, x_4) = \left(K_1, K_2, \frac{k_1 K_1}{k_3}, \frac{k_2 K_2}{k_4} \right) \in \mathcal{P}_{(K_1, K_2)}$$

Note that in this example, precluding the network's capacity for multistationarity did not depend on the reaction-rate parameters.

Example 7. Consider the following mass-action network

$$X_1 \xrightarrow{k_1} X_1 + P_1, \qquad X_2 \xrightarrow{k_2} X_2 + P_2, \qquad P_1 \xrightarrow{k_3,} 0, \qquad P_2 \xrightarrow{k_4,} 0,$$

$$P_1 + P_2 \underset{k_6}{\overset{k_5}{\rightleftharpoons}} P_1 P_2, \qquad X_2 + P_1 P_2 \underset{k_8}{\overset{k_7}{\rightleftharpoons}} X_2 P_1 P_2, \qquad X_2 P_1 P_2 \xrightarrow{k_9} X_2 P_1 P_2 + P_2.$$

Note that the first four reaction of this network are identical to those from our initial example (1) but with an additional three (two of which are reversible) biochemical processes considered. First, a heterodimerization of P_2 and P_2 occurs, which then binds to the X_2 promoter, and a P_2 is then produced from the bound gene. For simplicity we relabel $X_3 = P_2$, $X_4 = P_2$, $X_5 = P_1 P_2$, $X_6 = X_2 P_1 P_2$. The stoichiometric subspace is then given by

$$S = \mathrm{span}\left(\begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 2 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 1 \\ -1 \end{pmatrix} \right).$$

A simple computation shows that

$$W = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 \end{pmatrix} \in \mathbb{R}^{2 \times 6}$$

is a row-reduced matrix whose rows generate $S^{\perp}$. This means that $(1,0,0,0,0)^{\perp}$ and $(0,1,0,0,1)^{\perp}$ are the two generating conservation relations for the network. In particular, the values $x_1 = K_1$ and $x_2 + x_6 = K_2$ are constant for any trajectory. Note that the network is not conservative, however it is dissipative. The computation needed to show this is somewhat arduous but not terribly difficult, however we choose to omit it here. The difference between this computation and that of Example 1 is that here we actually seem to need the values $w_1, \dots, w_6 \in \mathbb{R}_{>0}$ to satisfy certain relations.

We now wish to preclude the existence of boundary steady states for the system, and begin by computing the siphons of the network. Since not all of the connected components are strongly connected a, it follows from Theorem 1 that the relevant ideal is given by

$$\mathfrak{m}_{1,G} = \langle X_1 \cdot (X_1X_3 - X_1), X_2 \cdot (X_2X_4 - X_2), X_3, X_4, X_3X_4 \cdot (X_5 - X3X_4), X_5 \cdot (X_3X_4 - X_5), \\ X_2X_5 \cdot (X_6 - X_2X_5), X_6 \cdot (X_2X_5 - X_6), X_6 \cdot (X_4X_6 - X_6)\rangle.$$

While this ideal might seem a lot more complicated than the one from example 1, a simple computation (again, using `Macaulay2`) reveals that the unique minimal prime of $\mathfrak{m}_{1,G}$ is given by $\langle X_1, \dots, X_6\rangle$, which once again means that the unique minimal siphon of the network is just the full set of species $\mathcal{X} = \{X_1, \dots, X_6\}$. Since this is trivially nonrelevant, it follows from Proposition 4 that the system does not contain any boundary steady states in any S-class $\mathcal{P}_c$ with $\mathcal{P}_c^+ \neq \emptyset$.

Now, let $c \in \mathbb{R}^2$ be such that $\mathcal{P}_c^+ \neq \emptyset$, and note that the function φ_c constructed from f and W is given by

$$\varphi_c(x) = \begin{pmatrix} x_1 - c_1 \\ x_2 + x_6 - c_2 \\ k_1x_1 - k_3x_3 - k_5x_3x_4 + k_6x_5 \\ k_2x_2 - k_4x_4 - k_5x_3x_4 + k_6x_5 + k_9x_6 \\ k_5x_3x_4 - k_6x_5 - k_7x_2x_5 + k_8x_6 \\ k_7x_2x_5 - k_8x_6 \end{pmatrix}.$$

The Jacobian matrix $M(x) = J_{\varphi_c}(x)$ is

$$M(x) = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ k_1 & 0 & -k_5x_4 - k_5 & -k_5x_3 & k_6 & 0 \\ 0 & k_2 & -k_5x_4 & -k_5x_3 - k_4 & k_6 & k_9 \\ 0 & -k_7x_5 & k_5x_4 & k_5x_3 & -k_7x_2 - k_6 & k_8 \\ 0 & k_7x_5 & 0 & 0 & k_7x_2 & -k_8 \end{pmatrix}.$$

The determinant of this matrix is

$$k_2k_3k_5k_7x_2x_3 - k_3k_5k_7k_9x_2x_3 + k_3k_4k_6k_7x_5 + k_3k_4k_6k_8.$$

The determinant has terms of sign $(-1)^3 = -1$, which means that we cannot immediately apply Theorem 7. Additionally, it is not possible to find a positive parametrization of the set $V_{>0}(f)$, which means that we unfortunately are not able to apply Corollary 3. This is due to the fact that the set of species $\{X_3, X_4, X_5, X_6\}$ is interacting (i.e. at least two of the species appear in a reactant together).

Example 8. Consider the following mass-action network

$$0 \xrightarrow{k_1} X_1 \xrightarrow{k_2} 0, \qquad X_1 + X_3 \xrightarrow{k_3} X_4 \xrightarrow{k_4} X_2X_3,$$

$$X_2 \xrightarrow{k_5} 0, \qquad X_1X_4 \underset{k_7}{\overset{k_6}{\rightleftharpoons}} X_5.$$

The stoichiometric subspace is given by

$$S = \operatorname{span}\left(\begin{pmatrix} 1\\0\\0\\0\\0 \end{pmatrix}, \begin{pmatrix} -1\\0\\-1\\1\\0 \end{pmatrix}, \begin{pmatrix} 0\\1\\1\\-1\\0 \end{pmatrix}, \begin{pmatrix} 1\\0\\0\\1\\-1 \end{pmatrix} \right).$$

A simple computation shows that

$$W = \begin{pmatrix} 0 & 0 & 1 & 1 & 1 \end{pmatrix} \in \mathbb{R}^{1\times 5}$$

is a row-reduced matrix whose rows generate $S^{\perp}$. This means that $(0,0,1,1,1)^{\perp}$ is the unique generating conservation relation for the network, and that for any trajectory of the system the value $x_3 + x_4 + x_5 = K_1$ is constant. Note that the network is not conservative, however it is dissipative, which once again is seen through choosing a suitable $w \in \mathbb{R}^5_{>0}$ and applying Theorem 3.

Now, in order to preclude the existence of boundary steady states we first note that the connected components of the network are not strongly connected, which means that the relevant ideal from Theorem 1 is given by

$$\mathfrak{m}_{1,G} = \langle X_1, X_1X_3\cdot(X_4 - X_1X_3), X_4\cdot(X_2X_3 - X_4), X_2, X_1X_4\cdot(X_5 - X_1X_4), X_5\cdot(X_1X_4 - X_5)\rangle.$$

Using `Macaulay2` we see that the unique minimal prime of $\mathfrak{m}_{1,G}$ is given by $\langle X_1, X_2, X_4, X_5\rangle$, which means that the unique minimal siphon of the network is $\{X_1, X_2, X_4, X_5\} \subseteq \mathcal{X}$. Note that this siphon does not contain the support of our unique generating conservation law, so it is not nonrelevant. This means that we are not able to apply Proposition 4 to preclude the existence of boundary steady states. However, a direct calculation shows that any steady state with $x_i = 0$ for any $i = 1, \ldots 5$ in some $\mathcal{P}_c$ with $\mathcal{P}_c^+ \neq \emptyset$ leads directly to a contradiction. We are therefore still able to preclude the existence of boundary steady states, although not as systematically as in our previous examples.

Now, let $c \in \mathbb{R}$ be such that $\mathcal{P}_c^+ \neq \emptyset$, and note that the function φ_c constructed from the formation-rate function f and W is given by

$$\varphi_c(x) = \begin{pmatrix} k_1 - k_2x_1 - k_3x_1x_3 - k_6x_1x_4 + k_7x_5 \\ k_4x_4 - k_5x_2 \\ x_3 + x_4 + x_5 - c_1 \\ k_3x_1x_3 - k_4x_4 - k_6x_1x_4 + k_7x_5 \\ k_6x_1x_4 - k_7x_5 \end{pmatrix}.$$

The Jacobian matrix $M(x) = J_{\varphi_c}(x)$ is

$$M(x) = \begin{pmatrix} -k_3x_3 - k_6x_4 - k_2 & 0 & -k_3x_1 & -k_6x_1 & k_7 \\ 0 & -k_5 & 0 & k_4 & 0 \\ 0 & 0 & 1 & 1 & 1 \\ k_3x_3 - k_6x_4 & 0k_3x_1 & -k_6x_1 - k_4 & k_7 & \\ k_6x_4 & 0 & 0 & k_6x_1 & -k_7 \end{pmatrix}$$

The determinant is a large polynomial which we omit. However, it contains terms of sign $(-1)^2 = 1$, which means that we cannot immediately apply Theorem 7. Instead, solving $f_1(x) = f_2(x) = f_4(x) = f_5(x)$ yields the following positive parametrization

$$x_1 = \frac{k_1}{k_3x_3 + k_2}, \qquad x_2 = \frac{k_1k_3x_3}{(k_3x_3 + k_2)k_5}, \qquad x_4 = \frac{k_1k_3x_3}{(k_3x_3 + k_2)k_4}, \qquad x_5 = \frac{k_1^2k_3k_6x_3}{(k_3x_3 + k_2)^2k_4k_7}.$$

Using this parametrization in the setting of Corollary 3, we see that the function $a(x_3)$ is a large rational function with positive denominator. Therefore, only the numerator of the function will be of interest. This numerator function $n(x)$ is a large polynomial function of the form $ax^3 + bx^2 + cx + d$, which we shall omit writing explicitly. The coefficients a, b, d are positive linear combinations of the parameters $k_1, \ldots, k_7 \in \mathbb{R}_{>0}$, while the linear term c has the form

$$c = k_1 k_2 k - 3^2 k_7 + 3k_2^2 k_3 k_4 k_7 - k_1^2 k_3^2 k_6,$$

Note that $n(x)$ attains negative values for $x \geq 0$ only if $c < 0$, which occurs if and only if

$$3k_2 k_4 k_7 < k_1 k_3 (k_1 k_6 - k_2 k_7)$$

Additionally $n(0) = d > 0$, which by the intermediate value theorem means that $n(x)$ attains negative values for $x \geq 0$ if and only if it has two distinct positive roots (and hence all three roots must be real). Finding explicit conditions to impose on the values $a, b, c, d \in \mathbb{R}$ (which themselves are functions of the parameters $k_1, \ldots, k_7 \in \mathbb{R}_{>0}$) in order for this to occur seems to be a difficult problem. Using the cubic discriminant (see for instance page 612 in [4]) we are able to get an explicit inequality that is equivalent to the existence of three distinct roots, and two positive of these roots being positive is equivalent to the inequality

$$\frac{-b + \sqrt{b^2 - 3ac}}{3a} > 0$$

We omit writing these inequalities explicitly since they are very long polynomial inequalities in $k_1, \ldots, k_7 \in \mathbb{R}_{>0}$ and they do not immediately yield any huge insight. However, it is indeed possible to give an explicit parameter region for when multistationary occurs. Note that if $c > 0$, that is if

$$3k_2 k_4 k_7 > k_1 k_3 (k_1 k_6 - k_2 k_7),$$

Then the numerator function will not attain any negative values, and so in this case there is a unique, non-degenerate steady state for the system.

Example 4, continued. Recall that we have shown this network to be both conservative and devoid of any boundary steady state in any S-class $\mathcal{P}_c$ with $\mathcal{P}_c^+ \neq \emptyset$. Now, for any $c \in \mathbb{R}^2$ be such that $\mathcal{P}_c^+ \neq \emptyset$ we have that the function φ_c constructed from f and W is given by

$$\varphi_c(x) = \begin{pmatrix} x_1 + x_2 + x_4 - c_1 \\ k_2 x_1 - k_4 x_2 \\ x_3 - x_4 - c_2 \\ -k_1 x_3 x_4 + k_3 x_1 + k_4 x_2 \end{pmatrix}.$$

The Jacobian matrix $M(x) = J_{\varphi_c}(x)$ is

$$M(x) = \begin{pmatrix} 1 & 1 & 0 & 1 \\ k_2 & -k_4 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ k_3 & k_4 & -k_1 x_4 & -k_1 x_3 \end{pmatrix}.$$

The determinant of this matrix is

$$(x_3 + x_4)(k_1 k_4 + k_1 k_2) + k_4(k_2 + k_3) > 0,$$

so from Theorem 7 we conclude that there is a unique, non-degenerate steady state for the system.

Example 9. Consider the following mass-action network

$$E + S \underset{k_2}{\overset{k_1}{\rightleftharpoons}} X_1 \xrightarrow{k_3} E + P, \qquad F + P \underset{k_5}{\overset{k_4}{\rightleftharpoons}} X_2 \xrightarrow{k_6} F + S,$$

which is based on Michaelis-Menten mechanisms of two enzymes E and P, a substrate S, a product P, and intermediate species X_1, X_2. We relabel $X_3 = E$, $X_4 = F$, $X_5 = S$, $X_6 = P$. The stoichiometric subspace is then given by

$$S = \text{Span}\left(\begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ -1 \\ 0 \\ -1 \end{pmatrix} \right).$$

A simple computation shows that

$$W = \begin{pmatrix} 1 & 0 & 0 & -1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & -1 & -1 \end{pmatrix} \in \mathbb{R}^{3\times 6}$$

is a row-reduced matrix whose rows generate $S^\perp$. This gives rise to the generating conservation laws $(1, 0, 0, -1, 1, 1)^\perp$, $(0, 1, 0, 1, 0, 0)^\perp$, and $(0, 0, 1, 1, -1, -1)^\perp$. In particular we see that

$$2\begin{pmatrix} 1 \\ 0 \\ 0 \\ -1 \\ 1 \\ 1 \end{pmatrix} + 2\begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \\ -1 \\ -1 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

is a positive conservation law, which means that the network is conservative. Note that the values $x_1 - x_4 + x_5 + x_6 = K_1$, $x_2 + x_4 = K_2$, and $x_3 + x_4 - x_5 - x_6 = K_3$ are therefore constant for any trajectory of the system.

We now wish to preclude the existence of boundary steady states for the system, and begin by computing the siphons of the network. Since not all of the connected components are strongly connected, theorem 1 tells us that the relevant ideal is given by

$$\begin{aligned} \mathfrak{m}_{1,G} = \langle X_3X_5 \cdot (X_1 - X_3X_5), &X_1 \cdot (X_3X_5 - X_1), X_1 \cdot (X_3X_6 - X_1), \\ &X_4X_6 \cdot (X_2 - X_4X_6), X_2 \cdot (X_4X_6 - X_2), X_2 \cdot (X_4X_5)\rangle. \end{aligned}$$

We compute the associated minimal primes using `Macaulay2`, and apply Theorem 1 to see that the minimal siphons of the network are given by

$$\{X_1, X_2, X_3, X_4\}, \qquad \{X_1, X_2, X_4, X_5\}, \qquad \{X_1, X_2, X_3, X_6\}.$$

Now, note that all three siphons are nonrelevant, since they respectively contain the support of the positive conservation laws $(1, 1, 1, 1, 0, 0)^\perp$, $(0, 1, 0, 1, 0, 0)^\perp$, and $(1, 0, 1, 0, 0, 0)^\perp$. From Corollary 4 we therefore conclude that the system does not contain any boundary steady states in any S-class $\mathcal{P}_c$ with $\mathcal{P}_c^+ \neq \emptyset$.

Now, let $c \in \mathbb{R}^3$ be such that $\mathcal{P}_c^+ \neq \emptyset$, and note that the function φ_c constructed from f and W is given by

$$\varphi_c(x) = \begin{pmatrix} x_1 - x_4 + x_5 + x_6 - c_1 \\ x_2 + x_4 - c_2 \\ x_3 + x_4 - x_5 - x_6 - c_3 \\ -k_4x_4x_6 + (k_5 + k_6)x_2 \\ -k_1x_3x_5 + k_2x_1 + k_6x_2 \\ k_3x_1 - k_4x_4x_6 + k_5x_2 \end{pmatrix}.$$

The Jacobian matrix $M(x) = J_{\varphi_c}(x)$ is

$$M = \begin{pmatrix} 1 & 0 & 0 & -1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & -1 & -1 \\ 0 & k_5 + k_6 & 0 & -k_4 x_6 & 0 & -k_4 x_4 \\ k_2 & k_6 & -k_1 x_5 & 0 & -k_1 x_3 & 0 \\ k_3 & k_5 & 0 & -k_4 x_6 & 0 & -k_4 x_4 \end{pmatrix},$$

and its determinant is

$$-k_1k_3k_4(x_3x_4 + x_3x_6) - k_1k_4k_6(x_3x_4 + x_4x_5) - k_1k_3k_5x_3 - k_1k_3k_6x_3 - k_2k_4k_6x_4 - k_3k_4k_6x_4.$$

Since all terms of the determinant are of sign $(-1)^3 = -1$ we conclude directly from Theorem 7 that the network has a unique steady state in $\mathcal{P}_c$, which must be non-degenerate.

6.3 Final Remarks

Throughout the project we have introduced and used a number of interesting tools for analysing reactions networks G and the associated mass-actions systems. Interestingly, the network structure alone will often play a large role in the dynamics of the associated system. Conversely, for the degree-theoretic method we have seen that the parameters $k_1, \ldots, k_r \in \mathbb{R}_{>0}$ of the system have the final say in whether the system exhibits multistationarity or not. A lot of things has to go right for us to be able to apply all our methods to a single specific network. The network often has to be conservative, the siphons must be nonrelevant, and we often have to be able to find a suitable positive parametrization in order to give specific criteria for whether the system is multistationary or not. However, the degree-theoretic method in particular is able to yield some very interesting and nontrivial results when applied to suitable networks.

References

[1] M. Atiyah and I. MacDonald. *Introduction to Commutative Algebra.* Addison-Wesley-Longman, 1969.

[2] A. Ben-Israel. Notes on linear inequalities, i: The intersection of the nonnegative orthant with complementary orthogonal subspaces. *Journal of Mathematical Analysis and Applications*, 1964.

[3] C. Conradi, E. Feliu, M. Mincheva, and C. Wiuf. Identifying parameter regions for multistationarity. *PLoS Computational Biology*, 2017.

[4] D. S. Dummit and R. M. Foote. *Abstract Algebra.* John Wiley and Sons, inc., third edition edition, 2004.

[5] E. Feliu and A. Dickenstein. *Algebraic Methods for Biochemical Reaction Networks, pages 5-29.* in progress.

[6] D. Grayson and M. Stillman. Macaulay 2, a software system for research in algebraic geometry. http://www2.macaulay2.com/Macaulay2/.

[7] J. R. Munkres. *Topology.* Upper Saddle River, NJ : Prentice Hall, Inc., 2 edition, 2000.

[8] A. Shiu and B. Sturmfels. Siphons in chemical reaction networks. *Bulletin of Mathemtical Biology*, 2010.

[9] E. D. Sontag. Structure and stability of certain chemical networks and applications to the kinetic proofreading model of t-cell receptor signal transduction. *IEEE. Transactions on Automatic Control*, 2000.

[10] R. Vandervorst. Topological methods for differential equations, *pages 1-30*, 2014.