On the bijectivity of families of exponential/generalized polynomial maps

Stefan Müller, Josef Hofbauer, Georg Regensburger

May 8, 2019

Abstract

We start from a parametrized system of d generalized polynomial equations (with real exponents) for d positive variables, involving n generalized monomials with n positive parameters. Existence and uniqueness of a solution for all parameters and for all right-hand sides is equivalent to the bijectivity of (every element of) a family of generalized polynomial/exponential maps. We characterize the bijectivity of the family of exponential maps in terms of two linear subspaces arising from the coefficient and exponent matrices, respectively. In particular, we obtain conditions in terms of sign vectors of the two subspaces and a nondegeneracy condition involving the exponent subspace itself. Thereby, all criteria can be checked effectively. Moreover, we characterize when the existence of a unique solution is robust with respect to small perturbations of the exponents or/and the coefficients. In particular, we obtain conditions in terms of sign vectors of the linear subspaces or, alternatively, in terms of maximal minors of the coefficient and exponent matrices. Finally, we present applications to chemical reaction networks with (generalized) mass-action kinetics.

Keywords: global invertibility, Hadamard's theorem, Descartes' rule, sign vectors, oriented matroids, perturbations, robustness, deficiency zero theorem

AMS subject classification: 12D10, 26C10, 52B99, 52C40

1 Introduction

Given two matrices $W = (w^1, \ldots, w^n)$, $\tilde{W} = (\tilde{w}^1, \ldots, \tilde{w}^n) \in \mathbb{R}^{d \times n}$ with $d \leq n$ and full rank, consider the parametrized system of generalized polynomial equations

$$\sum_{j=1}^{n} w_{ij}\, c_j\, x_1^{\tilde{w}_{1j}} \cdots x_d^{\tilde{w}_{dj}} = y_i, \quad i = 1, \ldots, d$$

Stefan Müller · Josef Hofbauer
Faculty of Mathematics, University of Vienna, Oskar-Morgenstern-Platz 1, 1090 Wien, Austria
Georg Regensburger
Institute for Algebra, Johannes Kepler University Linz, Altenberger Straße 69, 4040 Linz, Austria
Corresponding author: st.mueller@univie.ac.at

for d positive variables $x_i > 0$ (and right-hand sides y_i), involving the 'monomials' $c_j\, x_1^{\tilde{w}_{1j}} \cdots x_d^{\tilde{w}_{dj}} = c_j\, x^{\tilde{w}^j}$, $j = 1, \ldots, n$, in particular, the n positive parameters $c_j > 0$. In other words, $x \in \mathbb{R}^d_{>0}$, $y \in \mathbb{R}^d$, and $c \in \mathbb{R}^n_{>0}$. As in the theory of fewnomials [34, 46], the monomials are given, however, with a positive parameter associated to every monomial.

Writing the vector of monomials as $c \circ x^{\tilde{W}} \in \mathbb{R}^n_{>0}$, thereby introducing $x^{\tilde{W}} \in \mathbb{R}^n_{>0}$ as $(x^{\tilde{W}})_j = x^{\tilde{w}^j}$ and denoting componentwise multiplication by $\circ$, yields the compact form

$$W(c \circ x^{\tilde{W}}) = y.$$

Note that, for the existence of a positive solution x, the right-hand side y must lie in the interior of $C = \operatorname{cone} W$, the polyhedral cone generated by the columns of W. The question arises whether the above equation system has a unique positive solution $x \in \mathbb{R}^d_{>0}$, for all right-hand sides $y \in C^\circ \subseteq \mathbb{R}^d$ and all positive parameters $c \in \mathbb{R}^n_{>0}$. This question is equivalent to whether the generalized polynomial map $f_c \colon \mathbb{R}^d_{>0} \to C^\circ \subseteq \mathbb{R}^d$,

$$f_c(x) = W(c \circ x^{\tilde{W}})$$

or, equivalently, the exponential map $F_c \colon \mathbb{R}^d \to C^\circ \subseteq \mathbb{R}^d$,

$$F_c(x) = W(c \circ \mathrm{e}^{\tilde{W}^\mathsf{T} x})$$

is bijective for all $c \in \mathbb{R}^n_{>0}$.

In the context of chemical reaction networks (CRNs) with *generalized* mass-action kinetics [38, 39], the question is equivalent to whether every set of complex-balanced equilibria (an 'exponential manifold') intersects every stoichiometric class (an affine subspace) in exactly one point. For a motivation from CRNs, see Section 5 or [16]. The assumption of mass-action kinetics corresponds to $W = \tilde{W}$, and in this case there is indeed exactly one complex-balanced equilibrium in every stoichiometric class.

In case $W = \tilde{W}$, the map F_c also appears in toric geometry [20], where it is related to moment maps, and in statistics [41], where it is related to log-linear models. The following result (called Birch's Theorem in [48, 41, 13, 15, 24, 16]) guarantees the bijectivity of F_c for all $c > 0$.

Theorem 1 ([20], Section 4.2)**.** *Let $W = \tilde{W}$. Then the map F_c is a real analytic isomorphism of $\mathbb{R}^d$ onto C° for all $c > 0$.*

In this work, we characterize the simultaneous bijectivity of the maps F_c for all $c > 0$ (for given coefficients W and exponents $\tilde{W}$) in terms of (sign vectors of) the linear subspaces $S = \ker W \subseteq \mathbb{R}^n$ and $\tilde{S} = \ker \tilde{W} \subseteq \mathbb{R}^n$, see Theorem 14. Moreover, we characterize the robustness of bijectivity with respect to small perturbations of the exponents $\tilde{W}$ or/and the coefficients W, corresponding to small perturbations of the subspaces $\tilde{S}$ and S (in the Grassmannian), see Theorems 31, 40, 42.

Sufficient conditions for bijectivity have been given in previous work [38], using Brouwer degree, and parallel work [16], using differential topology. For a smaller class of maps [22], bijectivity has been proved, using Brouwer's fixed point theorem.

Our main technical tool is Hadamard's global inversion theorem which essentially states that a C^1-map is a diffeomorphism if and only if it is locally invertible and proper. By previous results [15, 38], the map F_c is locally invertible for all $c > 0$ if and only if it is injective for all $c > 0$ which can be characterized in terms of maximal minors of W and $\tilde{W}$ or, equivalently, in terms of sign vectors of the subspaces S and $\tilde{S}$, see Subsection 2.1. Most importantly, we show that F_c is proper if and only if it is 'proper along rays' and that properness for all $c > 0$ can be characterized in terms of sign vectors of S and $\tilde{S}$, together with a nondegeneracy condition depending on the subspace $\tilde{S}$ itself.

The crucial role of sign vectors in the characterization of existence and uniqueness of positive solutions to parametrized polynomial equations suggests a comparison with Descartes' rule of signs for univariate (generalized) polynomials [47, 35, 29]. A sharp rule [1] states that a univariate polynomial with given sign sequence has exactly one positive solution for all (positive) coefficients if and only if there is exactly one sign change. Indeed, this statement follows from our main result which can be seen as a multivariate generalization of the sharp Descartes' rule for exactly one positive solution.

Organization of the work and main results

In Section 2, we introduce the family of exponential maps F_c with $c > 0$ and discuss previous results on injectivity.

In Section 3, we present our main result, Theorem 14, characterizing the simultaneous bijectivity of the maps F_c, and the crucial Lemmas 11 and 16, regarding the properness of F_c. In Subsection 3.1, we discuss two extreme cases regarding the geometry of the cone C, namely, $C = \mathbb{R}^d$ and C is pointed. In Subsection 3.2, we show that the simultaneous bijectivity of the maps F_c cannot be characterized in terms of sign vectors only, cf. Example 20. Still, there are sufficient conditions for bijectivity in terms of sign vectors or in terms of faces of the Newton polytope, cf. Propositions 21 and 22.

In Section 4, we study the robustness of simultaneous bijectivity. In Subsection 4.1, we consider perturbations of the exponents $\tilde{W}$ and show that robustness of bijectivity is equivalent to robustness of injectivity which can be characterized in terms of sign vectors, cf. Theorem 31. The criterion involves the closure of a set of sign vectors and represents a simple sufficient condition for bijectivity, cf. Proposition 29. Equivalently, robustness can be characterized in terms of maximal minors. In Subsection 4.2, we consider perturbations of the coefficients W and characterize robustness of bijectivity again in terms of sign vectors (including another closure condition), cf. Theorem 40. In particular, robustness of bijectivity implies that either $C = \mathbb{R}^d$ or C is pointed. Finally, in Subsection 4.3, we consider general perturbations (of both exponents and coefficients) and characterize robustness of bijectivity in terms of sign vectors and maximal minors, cf. Theorem 42.

In Section 5, we present a derivation of our main problem from chemical reaction networks and applications of our main results. In particular, we formulate a deficiency zero theorem for generalized mass-action kinetics and a robust deficiency zero theorem for (generalized) mass-action kinetics, cf. Theorems 45 and 46.

Finally, we provide appendices on (A) oriented matroids and (B) a theorem of the alternative.

Notation

We denote the positive real numbers by $\mathbb{R}_{>0}$ and the nonnegative real numbers by $\mathbb{R}_{\geq 0}$. We write $x > 0$ for $x \in \mathbb{R}^n_{>0}$ and $x \geq 0$ for $x \in \mathbb{R}^n_{\geq 0}$. For vectors $x, y \in \mathbb{R}^n$, we denote their scalar product by $x \cdot y$ and their componentwise (Hadamard) product by $x \circ y$.

For a vector $x \in \mathbb{R}^n$, we obtain the sign vector $\operatorname{sign}(x) \in \{-, 0, +\}^n$ by applying the sign function componentwise, and we write

$$\operatorname{sign}(S) = \{\operatorname{sign}(x) \mid x \in S\}$$

for a subset $S \subseteq \mathbb{R}^n$.

For a vector $x \in F^n$ with $F = \mathbb{R}$ or $F = \{-, 0, +\}$, we denote its support by $\operatorname{supp}(x) = \{i \mid x_i \neq 0\}$. For a subset $X \subseteq F^n$, we say that a nonzero vector $x \in X$ has (inclusion-)minimal support, if $\operatorname{supp}(x') \subseteq \operatorname{supp}(x)$ implies $\operatorname{supp}(x') = \operatorname{supp}(x)$ for all nonzero $x' \in X$.

For a sign vector $\tau \in \{-, 0, +\}^n$, we introduce

$$\tau^- = \{i \mid \tau_i = -\}, \quad \tau^0 = \{i \mid \tau_i = 0\}, \quad \text{and} \quad \tau^+ = \{i \mid \tau_i = +\}.$$

In particular, $\operatorname{supp}(\tau) = \tau^- \cup \tau^+$. For a subset $T \subseteq \{-, 0, +\}^n$, we write

$$T_\oplus = T \cap \{0, +\}^n.$$

The inequalities $0 < -$ and $0 < +$ induce a partial order on $\{-, 0, +\}^n$: for sign vectors $\tau, \rho \in \{-, 0, +\}^n$, we write $\tau \leq \rho$ if the inequality holds componentwise. The product on $\{-, 0, +\}$ is defined in the obvious way. For $\tau, \rho \in \{-, 0, +\}^n$, we write $\tau \cdot \rho = 0$ (τ and ρ are orthogonal) if either $\tau_i \rho_i = 0$ for all i or there exist i, j with $\tau_i \rho_i = -$ and $\tau_j \rho_j = +$. For $T \subseteq \{-, 0, +\}^n$, we introduce the orthogonal complement

$$T^\perp = \{\tau \in \{-, 0, +\}^n \mid \tau \cdot \rho = 0 \text{ for all } \rho \in T\}.$$

Moreover, for $\tau, \rho \in \{-, 0, +\}^n$, we define the composition $\tau \circ \rho \in \{-, 0, +\}^n$ as $(\tau \circ \rho)_i = \tau_i$ if $\tau_i \neq 0$ and $(\tau \circ \rho)_i = \rho_i$ otherwise.

For a matrix $W \in \mathbb{R}^{d \times n}$, we denote its column vectors by $w^1, \ldots, w^n \in \mathbb{R}^d$. For any natural number n, we define $[n] = \{1, \ldots, n\}$. For $W \in \mathbb{R}^{d \times n}$ with $d \leq n$ and $I \subseteq [n]$ of cardinality d, we denote the square submatrix of W with column indices in I by W_I.

2 Families of exponential maps

Let $W \in \mathbb{R}^{d \times n}$, $\tilde{W} \in \mathbb{R}^{\tilde{d} \times n}$ be matrices with $d, \tilde{d} \leq n$ and full rank. Further, let

$$C = \operatorname{cone} W \subseteq \mathbb{R}^d$$

be the cone generated by the columns of W. Since W has full rank, the cone C has nonempty interior C°. Finally, let $c > 0$. We define the exponential map

$$
\begin{aligned}
F_c \colon \mathbb{R}^{\tilde{d}} &\to C^\circ \subseteq \mathbb{R}^d \\
x &\mapsto W(c \circ \mathrm{e}^{\tilde{W}^\mathsf{T} x}) = \sum_{i=1}^{n} c_i \, \mathrm{e}^{\tilde{w}^i \cdot x} \, w^i
\end{aligned}
\tag{1}
$$

and the related subspaces

$$
S = \ker W \subseteq \mathbb{R}^n \quad \text{and} \quad \tilde{S} = \ker \tilde{W} \subseteq \mathbb{R}^n. \tag{2}
$$

Note that injectivity and surjectivity of F_c only depend on S and $\tilde{S}$. In fact, let $V \in \mathbb{R}^{d \times n}$, $\tilde{V} \in \mathbb{R}^{\tilde{d} \times n}$ be such that $\ker V = S$, $\ker \tilde{V} = \tilde{S}$, and let

$$
G_c(x) = V(c \circ \mathrm{e}^{\tilde{V}^\mathsf{T} x})
$$

be the corresponding exponential map. Then $V = UW$, $\tilde{V} = \tilde{U}\tilde{W}$ for invertible matrices $U \in \mathbb{R}^{d \times d}$, $\tilde{U} \in \mathbb{R}^{\tilde{d} \times \tilde{d}}$, and

$$
G_c(x) = U F_c(\tilde{U}^\mathsf{T} x).
$$

2.1 Previous results on injectivity

In the context of multiple equilibria in mass-action systems [14] and geometric modeling [15], where $d = \tilde{d}$, it was shown that the map F_c is injective for all $c > 0$ if and only if F_c is a local diffeomorphism for all $c > 0$.

Theorem 2 (Theorem 7 and Corollary 8 in [15])**.** *Let F_c be as in* (1) *with $d = \tilde{d}$. Then the following statements are equivalent:*

1. *F_c is injective for all $c > 0$.*
2. *$\det(\frac{\partial F_c}{\partial x}) \neq 0$ for all x and all $c > 0$.*
3. *$\det(W_I) \det(\tilde{W}_I) \geq 0$ for all subsets $I \subseteq [n]$ of cardinality d (or '≤ 0' for all I) and $\det(W_I) \det(\tilde{W}_I) \neq 0$ for some I.*

In [38], we gave an alternative proof of this result and extended it to the case $d \neq \tilde{d}$, by using the sign vectors of the subspaces S and $\tilde{S}$.

Theorem 3 (Theorem 3.6 in [38])**.** *Let F_c be as in* (1) *and $S, \tilde{S}$ be as in* (2)*. Then the following statements are equivalent:*

1. *F_c is injective for all $c > 0$.*
2. *F_c is an immersion for all $c > 0$.*
   ($\frac{\partial F_c}{\partial x}$ is injective for all x and all $c > 0$.)
3. *$\operatorname{sign}(S) \cap \operatorname{sign}(\tilde{S}^\perp) = \{0\}$.*

Theorems 2 and 3 characterize the simultaneous injectivity of F_c (with $d = \tilde{d}$) for all $c > 0$ equivalently in terms of maximal minors and sign vectors.

Corollary 4. *Let $S, \tilde{S}$ be subspaces of $\mathbb{R}^n$ of dimension $n-d$ (with $d \leq n$). For every $W, \tilde{W} \in \mathbb{R}^{d\times n}$ (with full rank d) such that $S = \ker W$ and $\tilde{S} = \ker \tilde{W}$, the following statements are equivalent.*

1. $\operatorname{sign}(S) \cap \operatorname{sign}(\tilde{S}^\perp) = \{0\}$.
2. $\det(W_I)\det(\tilde{W}_I) \geq 0$ *for all subsets* $I \subseteq [n]$ *of cardinality* d *(or '*≤ 0*' for all* I*) and* $\det(W_I)\det(\tilde{W}_I) \neq 0$ *for some* I.

In the language of oriented matroids, Corollary 4 relates *chirotopes* (signs of maximal minors of W and $\tilde{W}$) to *vectors* (sign vectors of $S = \ker W$ and $\tilde{S} = \ker \tilde{W}$), see also Appendix A. Thereby, the sign vector condition is symmetric with respect to S and $\tilde{S}$.

Corollary 5 (Corollary 3.8 in [38])**.** *Let $S, \tilde{S}$ be subspaces of $\mathbb{R}^n$ of equal dimension. Then*

$$\operatorname{sign}(S) \cap \operatorname{sign}(\tilde{S}^\perp) = \{0\} \quad \textit{if and only if} \quad \operatorname{sign}(\tilde{S}) \cap \operatorname{sign}(S^\perp) = \{0\}.$$

For a direct proof of Corollaries 4 and 5, see also [12].

In further works on injectivity of families of exponential/generalized polynomial maps, the coefficient and exponent matrices need not have full rank, and injectivity is studied on affine subspaces, see [23, 19, 37, 6].

3 Bijectivity

A necessary condition for the bijectivity of the map F_c is $d = \tilde{d}$. In the rest of the paper, we consider F_c as in (1) with $d = \tilde{d}$ and the related subspaces $S, \tilde{S}$ as in (2).

A first sufficient condition for the bijectivity of the map F_c for all $c > 0$ (in terms of sign vectors of S and $\tilde{S}$) was given in [38], thereby extending Theorem 1.

Theorem 6 (Proposition 3.9 in [38])**.** *If* $\operatorname{sign}(S) = \operatorname{sign}(\tilde{S})$ *and* $(+,\dots,+)^\mathsf{T} \in \operatorname{sign}(S^\perp)$*, then the map F_c is a real analytic isomorphism for all $c > 0$.*

As it will turn out, $\operatorname{sign}(S) = \operatorname{sign}(\tilde{S})$ is sufficient for bijectivity, and the technical condition $(+,\dots,+)^\mathsf{T} \in \operatorname{sign}(S^\perp)$ in [38] is not needed, cf. Corollary 15. We note that Theorems 2, 3, and 6 allowed a first multivariate generalization of Descartes' rule of signs for at most/exactly one positive solution, see [37].

In order to characterize the simultaneous bijectivity of the map F_c for all $c > 0$, we start with the following observation.

Proposition 7. *The following statements are equivalent.*

1. *F_c is bijective for all $c > 0$.*
2. *F_c is a diffeomorphism for all $c > 0$.*
3. *F_c is a real analytic isomorphism for all $c > 0$.*

Proof. Let F_c be bijective for all $c > 0$. In particular, it is injective, and $\det(\frac{\partial F_c}{\partial x}) \neq 0$ for all x and $c > 0$, by Theorems 2 or 3. Hence, F_c is a local diffeomorphism for all $c > 0$. Further, F_c is real analytic and hence a local real analytic isomorphism for all $c > 0$. $\square$

Most importantly, we will use Hadamard's global inversion theorem.

Theorem 8 ([26], Theorem A in [25]). *A C^1-map $F \colon \mathbb{R}^d \to \mathbb{R}^d$ is a diffeomorphism if and only if the Jacobian $\det(\frac{\partial F}{\partial x}) \neq 0$ for all $x \in \mathbb{R}^d$ and $|F(x)| \to \infty$ whenever $|x| \to \infty$.*

Obviously, we need a slightly more general version of this result which follows from Satz II in [5] or Theorem B in [25].

Theorem 9. *Let $U \subseteq \mathbb{R}^d$ be open and convex. A C^1-map $F \colon \mathbb{R}^d \to U$ is a diffeomorphism if and only if the Jacobian $\det(\frac{\partial F}{\partial x}) \neq 0$ for all $x \in \mathbb{R}^d$ and F is proper.*

Recall that a map F between two topological spaces is *proper*, if $F^{-1}(K)$ is compact for each compact subset K of the target space. This is obviously necessary for the inverse F^{-1} to be continuous.

Lemma 10. *Let $U \subseteq \mathbb{R}^d$ be open. A continuous map $F \colon \mathbb{R}^d \to U$ is proper if and only if, for sequences x_n in $\mathbb{R}^d$ with $|x_n| = 1$ and $x_n \to x$ and t_n in $\mathbb{R}_{>0}$ with $t_n \to \infty$, $F(x_n t_n) \to y$ implies $y \in \partial U$.*

Proof. Suppose F is proper and $F(x_n t_n) \to y$, but $y \in U$. Take a closed ball $K \subseteq U$ around y. Then $F^{-1}(K)$ contains the unbounded sequence $(x_n t_n)_{n \geq N}$ for some positive N and hence is not compact, a contradiction.

Conversely, let K be a compact subset of U. We need to show that every sequence X_n in $F^{-1}(K)$ has an accumulation point. Since $F^{-1}(K)$ is closed, we only need to show that X_n has a bounded subsequence. Suppose not, then $|X_n| \to \infty$. Since $F(X_n) \in K$, there is a subsequence (call it X_n again) such that $F(X_n) \to y \in K$. Now there is a subsubsequence (call it X_n again) such that $x_n = X_n/|X_n| \to x$, that is, the sequence x_n on the unit sphere converges. With $t_n = |X_n|$, we have $F(x_n t_n) \to y \in K \subset U$, a contradiction. $\square$

In particular, if F is proper, then, for all nonzero $x \in \mathbb{R}^d$, $F(xt) \to y$ as $t \to \infty$ implies $y \in \partial U$. That is, if the function values converge along a ray, then the limit lies on the boundary of the range.

By Lemma 11 below, the map F_c under consideration is proper, if it is 'proper along rays'. Before we prove this result, we discuss the behaviour of F_c along a ray. For $x \in \mathbb{R}^d$ and $\lambda \in \mathbb{R}$, we introduce

$$I_{x,\lambda} = \{i \mid \tilde{w}^i \cdot x = \lambda\}$$

and write

$$F_c(xt) = \sum_{\lambda} \sum_{i \in I_{x,\lambda}} c_i \, \mathrm{e}^{\lambda t} \, w^i,$$

where a sum over the empty set is defined as zero. For $x \in \mathbb{R}^d$ and $c > 0$, let $\lambda_{\max}$ be the largest λ such that $\sum_{i \in I_{x,\lambda}} c_i w^i \neq 0$. If $\lambda_{\max} > 0$, then

$$F_c(xt)\,\mathrm{e}^{-\lambda_{\max} t} \to \sum_{i \in I_{x,\lambda_{\max}}} c_i w^i \neq 0$$

as $t \to \infty$ and hence $|F_c(xt)| \to \infty$. If $\lambda_{\max} \leq 0$, then

$$F_c(xt) \to \sum_{i \in I_{x,0}} c_i w^i \in C$$

as $t \to \infty$. In this case, any vector w^i with $i \in I_{x,\lambda}$ and $\lambda > 0$ (and hence $\sum_{i \in I_{x,\lambda}} c_i w^i = 0$) lies in the lineality space of C, see Appendix A. If $\lambda_{\max} < 0$, then $F_c(xt) \to 0$. As a result, we have the following fact:

For every $x \in \mathbb{R}^d$, either $|F_c(xt)| \to \infty$ as $t \to \infty$ or $F_c(xt) \to y \in C$.

Lemma 11. *The map F_c is proper, if*

$$F_c(xt) \to y \quad \textit{as} \quad t \to \infty \quad \textit{implies} \quad y \in \partial C \tag{$*$}$$

for all nonzero $x \in \mathbb{R}^d$.

Proof. We assume that the ray condition $(*)$ holds for all nonzero $x \in \mathbb{R}^d$.

Let $x \in \mathbb{R}^d$ with $|x| = 1$. In order to apply Lemma 10, we consider sequences x_n in $\mathbb{R}^d$ with $|x_n| = 1$ and $x_n \to x$ and t_n in $\mathbb{R}_{>0}$ with $t_n \to \infty$.

To begin with, we show that $|F_c(xt)| \to \infty$ as $t \to \infty$ implies $|F_c(x_n t_n)| \to \infty$ as $n \to \infty$. Suppose $|F_c(xt)| \to \infty$, that is, there is $\lambda > 0$ such that $F_c(xt)\,\mathrm{e}^{-\lambda t} \to \sum_{i \in I_{x,\lambda}} c_i w^i \neq 0$ as $t \to \infty$. For x' close to x, we have the partition

$$I_{x,\lambda} = I_{x',\mu_1} \cup \dots \cup I_{x',\mu_p}$$

with μ_j close to λ and hence $\mu_j > \frac{\lambda}{2}$. Hence, there exists a largest μ_j such that $\sum_{i \in I_{x',\mu_j}} c_i w^i \neq 0$. Otherwise,

$$\sum_{i \in I_{x,\lambda}} c_i w^i = \sum_{i \in I_{x',\mu_1}} c_i w^i + \ldots + \sum_{i \in I_{x',\mu_p}} c_i w^i = 0.$$

Additionally, there may exist an even larger μ with $\sum_{i \in I_{x',\mu}} c_i w^i \neq 0$. In any case, there is $\lambda' > \frac{\lambda}{2}$ such that

$$F_c(x't)\,\mathrm{e}^{-\lambda' t} \to \sum_{i \in I_{x',\lambda'}} c_i w^i \neq 0$$

as $t \to \infty$ and hence $|F_c(x't)|\,\mathrm{e}^{-\frac{\lambda}{2}t} > \gamma$ with $\gamma > 0$ independent of x'; that is, $|F_c(x't)| > \gamma\,\mathrm{e}^{\frac{\lambda}{2}t}$ as $t \to \infty$. Hence $|F_c(x_n t_n)| > \gamma\,\mathrm{e}^{\frac{\lambda}{2}t_n}$ as $n \to \infty$; that is, $|F_c(x_n t_n)| \to \infty$, as claimed.

In case $C = \mathbb{R}^d$ ($\partial C = \emptyset$), the ray condition $(*)$ implies $|F_c(xt)| \to \infty$ as $t \to \infty$ and hence $|F_c(x_n t_n)| \to \infty$ as $n \to \infty$. By Lemma 10, F_c is proper.

In case $C \neq \mathbb{R}^d$, assume $F_c(x_n t_n) \to y'$ as $n \to \infty$. Then, $F_c(xt) \to y$ as $t \to \infty$, by the first argument in the proof and the fact before the lemma. In particular, $\sum_{i \in I_{x,\lambda}} c_i w^i = 0$ for $\lambda > 0$ and $y = \sum_{i \in I_{x,0}} c_i w^i$. Hence, vectors w^i with $i \in I_{x,\lambda}$ and $\lambda > 0$ lie in the lineality space of C. By the ray condition $(*)$, $y \in \partial C$, and hence

$$\operatorname{cone}(w^i \mid i \in I_{x,0}) \subseteq \partial C.$$

Finally, we write

$$F_c(x_n t_n) = \sum_{i=1}^{n} c_i \, \mathrm{e}^{\tilde{w}^i \cdot x_n \, t_n} \, w^i = \sum_{\lambda} \sum_{i \in I_{x,\lambda}} c_i \, \mathrm{e}^{\tilde{w}^i \cdot x_n \, t_n} \, w^i.$$

For x_n close to x, we have $\tilde{w}^i \cdot x_n$ close to λ for $i \in I_{x,\lambda}$, in particular, $\sum_{i \in I_{x,\lambda}} c_i \, \mathrm{e}^{\tilde{w}^i \cdot x_n \, t_n} \, w^i \to 0$ for $\lambda < 0$. The limit $F_c(x_n t_n) \to y'$ as $n \to \infty$ implies

$$\sum_{i \in I_{x,0}} c_i \, \mathrm{e}^{\tilde{w}^i \cdot x_n \, t_n} \, w^i + \sum_{\lambda > 0} \sum_{i \in I_{x,\lambda}} c_i \, \mathrm{e}^{\tilde{w}^i \cdot x_n \, t_n} \, w^i \to y',$$

and $y' \in \partial C$ since the sum of a vector in ∂C and a vector in the lineality space of C lies in ∂C. By Lemma 10, F_c is proper. □

Let $F_c(xt) \to y$ as $t \to \infty$ along the ray given by x and $F_c(x_n t_n) \to y'$ as $n \to \infty$ for a sequence $x_n t_n$ (with $x_n \to x$ and $t_n \to \infty$), approaching the ray. In the proof of Lemma 11, we have shown that, if $y = 0$, then $y' \in L$, where L is the lineality space of C. In general, if $y \in C_x = \operatorname{cone}(w^i \mid i \in I_{x,0})$, then $y' \in C_x + L$. Note that there are only finitely many index sets $I_{x,0}$ and hence finitely many limit points $y = \sum_{i \in I_{x,0}} c_i w^i$ (for fixed $c > 0$), whereas every $y' \in \partial C$ arises as a limit point (if F_c is surjective).

Using Theorem 9 (Hadamard's global inversion theorem) together with Theorems 2 or 3 and Lemma 11, we summarize our findings.

Corollary 12. *The map F_c is bijective for all $c > 0$ if and only if F_c is injective for all $c > 0$ and the ray condition $(*)$ in Lemma 11 holds for all nonzero $x \in \mathbb{R}^d$ and all $c > 0$.*

By Theorems 2 or 3, the simultaneous injectivity of F_c for all $c > 0$ can be characterized in terms of sign vectors of the subspaces S and $\tilde{S}$. By Lemma 16 below, the ray condition $(*)$ (for all nonzero $x \in \mathbb{R}^d$ and all $c > 0$) can be characterized in terms of sign vectors of S and $\tilde{S}$ together with a nondegeneracy condition depending on sign vectors of S and on the subspace $\tilde{S}$ itself.

Definition 13. Let $S, \tilde{S}$ be subspaces of $\mathbb{R}^n$. The pair $(S, \tilde{S})$ is called *nondegenerate* if, for every $z \in \tilde{S}^\perp$ with a positive component,

- there is $I = \{i \mid z_i = \lambda\}$ with $\lambda > 0$, defining $\pi \in \{0, +\}^n$ with $\pi^+ = I$, such that $\pi \notin \operatorname{sign}(S)_\oplus$ or
- for $\tilde{\tau} = \operatorname{sign}(z) \in \operatorname{sign}(\tilde{S}^\perp)$, there is a nonzero $\tau \in \operatorname{sign}(S^\perp)_\oplus$ such that $\tilde{\tau}^0 \subseteq \tau^0$.

As our main result, we obtain a characterization of the simultaneous bijectivity of F_c for all $c > 0$ in terms of the subspaces S and $\tilde{S}$.

Theorem 14. *The map F_c is a diffeomorphism for all $c > 0$ if and only if*

(i) $\operatorname{sign}(S) \cap \operatorname{sign}(\tilde{S}^\perp) = \{0\}$,

(ii) for every nonzero $\tilde{\tau} \in \operatorname{sign}(\tilde{S}^\perp)_\oplus$, there is a nonzero $\tau \in \operatorname{sign}(S^\perp)_\oplus$ such that $\tau \leq \tilde{\tau}$, and

(iii) the pair $(S, \tilde{S})$ is nondegenerate.

Theorem 14 immediately implies Theorems 1 and 6 ('Birch's Theorem' and its first extension).

Corollary 15. *The map F_c is a diffeomorphism for all $c > 0$ if* $\operatorname{sign}(S) = \operatorname{sign}(\tilde{S})$.

Proof. By Corollary 53 in Appendix B, $\operatorname{sign}(S^\perp) = \operatorname{sign}(S)^\perp$. Hence, $\operatorname{sign}(S) = \operatorname{sign}(\tilde{S})$ implies conditions (i) and (ii) in Theorem 14. Now, for $z \in \tilde{S}^\perp$ with a positive component $z_i = \lambda > 0$, consider $\pi \in \{0, +\}^n$ with $\pi^+ = \{i \mid z_i = \lambda\}$ and $\tilde{\tau} = \operatorname{sign}(z) \in \operatorname{sign}(\tilde{S}^\perp)$. Obviously, $\pi \cdot \tilde{\tau} \neq 0$ and hence $\pi \notin \operatorname{sign}(\tilde{S})_\oplus = \operatorname{sign}(S)_\oplus$. That is, $(S, \tilde{S})$ is nondegenerate, as required by condition (iii). □

We note that condition (i) in Theorem 14 can also be characterized in terms of maximal minors of the matrices W and $\tilde{W}$, cf. Corollary 4.

Condition (ii) can be reformulated using faces of the cones $C = \operatorname{cone} W$ and $\tilde{C} = \operatorname{cone} \tilde{W}$:

(ii) for every proper face $\tilde{f}$ of $\tilde{C}$ with $\tilde{I} = \{i \mid \tilde{w}^i \in \tilde{f}\}$, there is a proper face f of C with $I = \{i \mid w^i \in f\}$ such that $\tilde{I} \subseteq I$.

Indeed, a face f of C with $I = \{i \mid w^i \in f\}$ corresponds to a supporting hyperplane with normal vector x such that $w^i \cdot x = 0$ for $i \in I$ and $w^i \cdot x > 0$ otherwise (for w^i lying on the positive side of the hyperplane). Hence f is characterized by the nonnegative sign vector $\tau = \operatorname{sign}(W^\mathsf{T} x) \in \operatorname{sign}(S^\perp)_\oplus$ with $\tau^0 = I$. Analogously, a face $\tilde{f}$ of $\tilde{C}$ with $\tilde{I} = \{i \mid \tilde{w}^i \in \tilde{f}\}$ is characterized by a nonnegative sign vector $\tilde{\tau} \in \operatorname{sign}(\tilde{S}^\perp)_\oplus$ with $\tilde{\tau}^0 = \tilde{I}$. Clearly, $\tilde{I} \subseteq I$ is equivalent to $\tau \leq \tilde{\tau}$. (For more details on sign vectors and face lattices, see Appendix A.)

Condition (iii) concerns nondegeneracy. The second condition in Definition 13, on sign vectors $\tilde{\tau} = \operatorname{sign}(z) \in \operatorname{sign}(\tilde{S}^\perp)$, corresponds to condition (ii), on nonnegative sign vectors $\tilde{\tau} \in \operatorname{sign}(\tilde{S}^\perp)_\oplus$. The first condition on $z \in \tilde{S}^\perp$ can also be interpreted geometrically (in terms of the columns of $W, \tilde{W}$). Note that $\tilde{S}^\perp = (\ker \tilde{W})^\perp = \operatorname{im} \tilde{W}^\mathsf{T}$ and $z = \tilde{W}^\mathsf{T} x$ for some $x \in \mathbb{R}^d$. Hence, the set

$$I = \{i \mid z_i = \lambda\} = \{i \mid \tilde{w}^i \cdot x = \lambda\} = I_{x,\lambda}$$

with $\lambda > 0$ indicates equal positive components z_i or, geometrically, equal positive projections of columns $\tilde{w}^i$ (on x). The corresponding columns w^i must not be positively dependent, as expressed by the condition $\pi \notin \operatorname{sign}(S)_\oplus$ for the nonnegative sign vector $\pi \in \{0, +\}^n$ with $\pi^+ = I$.

It remains to prove Lemma 16.

Lemma 16. *The ray condition* $(*)$ *in Lemma 11 holds for all nonzero* $x \in \mathbb{R}^d$ *and for all* $c > 0$ *if and only if conditions (ii) and (iii) in Theorem 14 hold.*

Proof. For nonzero $x \in \mathbb{R}^d$, let $\lambda_x = \max_i \tilde{w}^i \cdot x$. We show the following two statements. Condition (ii) is equivalent to: the ray condition $(*)$ holds for all nonzero x with $\lambda_x \le 0$ and all $c > 0$. Condition (iii) is equivalent to: the ray condition $(*)$ holds for all nonzero x with $\lambda_x > 0$ and all $c > 0$.

(ii): If $\lambda_x \le 0$, then $\tilde{\tau} = \operatorname{sign}(-\tilde{W}^\mathsf{T} x) \in \operatorname{sign}(\tilde{S}^\perp)_\oplus$ defines a proper face of $\tilde{C}$ and $F_c(xt) \to \sum_{i \in \tilde{\tau}^0} c_i w^i$ as $t \to \infty$. The ray condition $(*)$ for all $c > 0$ is equivalent to $\sum_{i \in \tilde{\tau}^0} c_i w^i \in \partial C$ for all $c > 0$. That is, there is a proper face of C characterized by a nonzero $\tau \in \operatorname{sign}(S^\perp)_\oplus$ such that $\tilde{\tau}^0 \subseteq \tau^0$. Equivalently, $\tau \le \tilde{\tau}$, that is, (ii) for $\tilde{\tau}$.

By varying over all nonzero $x \in \mathbb{R}^d$ with $\lambda_x \le 0$, all nonzero $\tilde{\tau} \in \operatorname{sign}(\tilde{S}^\perp)_\oplus$ are covered.

(iii): If $\lambda_x > 0$, then $z = \tilde{W}^\mathsf{T} x \in \tilde{S}^\perp$ has a positive component. Using the fact before Lemma 11, the ray condition $(*)$ for all $c > 0$ is equivalent to

for all $c > 0$,

(α) either there is $\lambda > 0$ such that $F_c(xt)\,\mathrm{e}^{-\lambda t} \to \sum_{i \in I_{x,\lambda}} c_i w^i \neq 0$ as $t \to \infty$

(β) or $F_c(xt) \to \sum_{i \in I_{x,0}} c_i w^i \in \partial C$.

This is further equivalent to

(a) there is $\lambda > 0$ such that, for all $c > 0$, $\sum_{i \in I_{x,\lambda}} c_i w^i \neq 0$ or

(b) $\sum_{i \in I_{x,0}} c_i w^i \in \partial C$ for all $c > 0$.

To see this, note that the sets $I_{x,\lambda}$ are disjoint and the sums $\sum_{i \in I_{x,\lambda}} c_i w^i$ involve different coefficients c_i for different λ.
($\Rightarrow$): Assume $\neg$(a), that is, for all $\lambda > 0$, there exists $c > 0$ such that $\sum_{i \in I_{x,\lambda}} c_i w^i = 0$. Then, $\sum_{i \in I_{x,0}} c_i w^i \in \partial C$ for all $c > 0$, that is, (b).
($\Leftarrow$): Clearly, (a) implies (α) for all $c > 0$. Finally, assume (b) and let $c > 0$. Then, either (α) or, for all $\lambda > 0$, $\sum_{i \in I_{x,\lambda}} c_i w^i = 0$. In the latter case, $F_c(xt) \to \sum_{i \in I_{x,0}} c_i w^i$ with $\sum_{i \in I_{x,0}} c_i w^i \in \partial C$, that is, ($\beta$).

Finally, (a) or (b) is equivalent to

- there is $I_{x,\lambda} = \{i \mid z_i = \lambda\}$ with $\lambda > 0$ such that $c \notin \ker W = S$ for all $c \ge 0$ with $\operatorname{supp}(c) = I_{x,\lambda}$, that is, there is $\pi \in \{0,+\}^n$ with $\pi^+ = I_{x,\lambda}$ such that $\pi \notin \operatorname{sign}(S)_\oplus$, or
- for $\tilde{\tau} = \operatorname{sign}(z) \in \operatorname{sign}(\tilde{S}^\perp)$ and hence $\tilde{\tau}^0 = I_{x,0}$, there is a proper face of C, characterized by a nonzero $\tau \in \operatorname{sign}(S^\perp)_\oplus$, such that $\tilde{\tau}^0 \subseteq \tau^0$,

that is, (iii) for z.

By varying over all nonzero $x \in \mathbb{R}^d$ with $\lambda_x > 0$, all $z \in \tilde{S}^\perp$ with a positive component are covered. $\square$

3.1 Special cases: $C = \mathbb{R}^d$ or C is pointed

We discuss the conditions for bijectivity in Theorem 14 for two extreme cases, regarding the geometry of the cones $C = \operatorname{cone} W$ and $\tilde{C} = \operatorname{cone} \tilde{W}$.

If $C = \mathbb{R}^d$ (that is, $\operatorname{sign}(S^\perp)_\oplus = \{0\}$), then condition (ii) is equivalent to $\tilde{C} = \mathbb{R}^d$. Hence, if $C = \mathbb{R}^d$ and F_c is bijective for all $c > 0$, then $\tilde{C} = \mathbb{R}^d$. However, the converse does not hold.

Example 17. Let F_c be given by the matrices

$$\tilde{W} = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \end{pmatrix} \quad \text{and} \quad W = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}.$$

Then $\tilde{C} = \mathbb{R}^2$ and F_c is bijective for all $c > 0$. However, $C \neq \mathbb{R}^2$.

If $(+, \dots, +)^\mathsf{T} \in \operatorname{sign}(S^\perp)$ (that is, C is pointed and no column of W is zero), then condition (iii) holds (since $\operatorname{sign}(S)_\oplus = \{0\}$), and conditions (i) and (ii) imply $(+, \dots, +)^\mathsf{T} \in \operatorname{sign}(\tilde{S}^\perp)$ (by Proposition 19 below). Hence, if $(+, \dots, +)^\mathsf{T} \in \operatorname{sign}(S^\perp)$ and F_c is bijective for all $c > 0$, then $(+, \dots, +)^\mathsf{T} \in \operatorname{sign}(\tilde{S}^\perp)$. However, the converse does not hold.

Example 18. Let F_c be given by the matrices

$$\tilde{W} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \quad \text{and} \quad W = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}.$$

Then, $\tilde{C} = \mathbb{R}^2_{\geq 0}$, $(+, +, +)^\mathsf{T} \in \operatorname{sign}(\tilde{S}^\perp)$, and F_c is bijective for all $c > 0$. However, $C = \mathbb{R} \times \mathbb{R}_{\geq 0}$ and $(+, +, +)^\mathsf{T} \notin \operatorname{sign}(S^\perp)$.

If $(+, \dots, +)^\mathsf{T} \in \operatorname{sign}(S^\perp)$ (that is, C is pointed and no column of W is zero), then conditions (i) and (ii) imply the surjectivity of F_c for all $c > 0$ and, by the following result, $(+, \dots, +)^\mathsf{T} \in \operatorname{sign}(\tilde{S}^\perp)$.

Proposition 19. *Let* $(+, \dots, +)^\mathsf{T} \in \operatorname{sign}(S^\perp)$. *If* F_c *is surjective, then* $(+, \dots, +)^\mathsf{T} \in \operatorname{sign}(\tilde{S}^\perp)$.

Proof. By surjectivity, the image of F_c contains points arbitrarily close to zero. Hence, there is a sequence X_k in $\mathbb{R}^d$ such that $F_c(X_k) \to 0$ as $k \to \infty$. Since $(+, \dots, +)^\mathsf{T} \in \operatorname{sign}(S^\perp) = \operatorname{sign}(\operatorname{im} W^\mathsf{T})$, there is $y \in \mathbb{R}^d$ such that $y \cdot w^i > 0$ for all $i \in [n]$. Now,

$$y \cdot F_c(X_k) = \sum_{i=1}^{n} c_i \, (y \cdot w^i) \, \mathrm{e}^{\tilde{w}^i \cdot X_k}$$

is a sum of positive terms converging to zero, and hence each term goes to zero. This implies $\tilde{w}^i \cdot X_k < 0$ for large k, for all $i \in [n]$. Hence,

$$(+, \dots, +)^\mathsf{T} = \operatorname{sign}(-\tilde{W}^\mathsf{T} X_k) \in \operatorname{sign}(\operatorname{im} \tilde{W}^\mathsf{T})_\oplus = \operatorname{sign}(\tilde{S}^\perp)_\oplus.$$

□

3.2 Sign-vector conditions

In general, the simultaneous bijectivity of F_c for all $c > 0$, in particular, condition (iii) in Theorem 14, cannot be characterized in terms of sign vectors of S and $\tilde{S}$.

Example 20. Let F_c be given by the matrices

$$\tilde{W} = \begin{pmatrix} 1 & 1 & 0 & 0 & -1 & \tilde{w} \\ 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 \end{pmatrix} \quad \text{and} \quad W = \begin{pmatrix} 0 & 0 & 1 & 1 & -1 & 0 \\ 1 & -1 & 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & -1 & 0 & 0 \end{pmatrix},$$

involving the parameter $\tilde{w} > 0$. Obviously, $\tilde{C} = C = \mathbb{R}^3$. For $\tilde{w} = 1$ or $\tilde{w} \in [2, \infty)$, the map F_c is injective for all $c > 0$, but not bijective, whereas for $\tilde{w} \in (0, 1)$ or $\tilde{w} \in (1, 2)$, the map F_c is bijective for all $c > 0$. Clearly, the sign vectors $\operatorname{sign}(\tilde{S}) = \operatorname{sign}(\ker \tilde{W})$ do not depend on $\tilde{w}$ and hence cannot characterize bijectivity.

In general, condition (iii) depends on the subspace $\tilde{S}$ itself. Still,

- condition (iii) holds trivially if $(+, \dots, +)^\mathsf{T} \in \operatorname{sign}(S^\perp)$, see Section 3.1,
- there is a (weakest) condition (iv) in terms of sign vectors of S and $\tilde{S}$ *sufficient* for nondegeneracy, see Proposition 21, and
- there is a *sufficient* condition for nondegeneracy using faces of the Newton polytope $\tilde{P}$, see Proposition 22. (Thereby, faces of $\tilde{P}$ correspond to nonnegative sign vectors of an affine subspace related to $\tilde{S}$.)

Proposition 21. *Let $S, \tilde{S}$ be subspaces of $\mathbb{R}^n$. If*

(iv) for all $\tilde{\tau} \in \operatorname{sign}(\tilde{S}^\perp)$ with $\tilde{\tau}^+ \neq \emptyset$,

- *there is no $\pi \in \operatorname{sign}(S)_\oplus$ with $\pi^+ = \tilde{\tau}^+$*
- *or there is no $\rho \in \operatorname{sign}(S)$ with $\tilde{\tau}^+ \cup \tilde{\tau}^- \subseteq \rho^+$*

then the pair $(S, \tilde{S})$ is nondegenerate. That is, (iv) $\Rightarrow$ (iii).

Proof. Assume that $(S, \tilde{S})$ is degenerate, in particular, that $z \in \tilde{S}^\perp$ with a positive component violates nondegeneracy, and let $\tilde{\tau} = \operatorname{sign}(z) \in \operatorname{sign}(\tilde{S}^\perp)$, where $\tilde{\tau}^+ \neq \emptyset$.

For every index set $I = \{i \mid z_i = \lambda\}$ with $\lambda > 0$, the sign vector $\pi \in \{0, +\}^n$ with $\pi^+ = I$ satisfies $\pi \in \operatorname{sign}(S)_\oplus$. Clearly, the index sets I cover $\tilde{\tau}^+ = \{i \mid z_i > 0\}$ and, by composition, there is $\pi \in \operatorname{sign}(S)_\oplus$ with $\pi^+ = \tilde{\tau}^+$.

Further, there is no nonzero $\tau \in \operatorname{sign}(S^\perp)_\oplus$ such that $\tilde{\tau}^0 \subseteq \tau^0$, that is, $\tau \leq |\tilde{\tau}|$. Thereby, $|\tilde{\tau}| \in \{0, +\}^n$ with $|\tilde{\tau}|^0 = \tilde{\tau}^0$ and $|\tilde{\tau}|^+ = \tilde{\tau}^+ \cup \tilde{\tau}^-$. By Corollary 52 in Appendix B, there is $\rho \in \operatorname{sign}(S)$ such that $\rho \geq |\tilde{\tau}|$, that is, $|\tilde{\tau}|^+ \subseteq \rho^+$. □

Finally, we formulate a sufficient condition for nondegeneracy using faces of the Newton polytope $\tilde{P} = \operatorname{conv} \tilde{W}$, the convex hull of the columns of $\tilde{W}$. A face $\tilde{f}$ of $\tilde{P}$ with $\tilde{I} = \{i \mid \tilde{w}^i \in \tilde{f}\}$ corresponds to a supporting affine hyperplane with normal vector $x \in \mathbb{R}^d$ and $\lambda \in \mathbb{R}$ such that $\tilde{w}^i \cdot x = \lambda$ for $i \in \tilde{I}$ and $\tilde{w}^i \cdot x < \lambda$ otherwise; that is, $\tilde{I} = I_{x,\lambda}$. It further corresponds to $z = \tilde{W}^\mathsf{T} x \in \tilde{S}^\perp$, where $\tilde{I} = \{i \mid z_i = \lambda\}$. If $\lambda > 0$, we call the face $\tilde{f}$ of $\tilde{P}$ *positive*, and $z \in \tilde{S}^\perp$ has a positive component.

Proposition 22. *Let $S, \tilde{S}$ be subspaces of $\mathbb{R}^n$, $\tilde{W} \in \mathbb{R}^{d \times n}$ be a matrix with full rank such that $\ker \tilde{W} = \tilde{S}$, and $\tilde{P} = \operatorname{conv} \tilde{W}$ be the Newton polytope. The pair $(S, \tilde{S})$ is nondegenerate, if, for every positive face $\tilde{f}$ of $\tilde{P}$ with $\tilde{I} = \{i \mid \tilde{w}^i \in \tilde{f}\}$, the sign vector $\pi \in \{0,+\}^n$ with $\pi^+ = \tilde{I}$ satisfies $\pi \notin \operatorname{sign}(S)_\oplus$.*

Proof. Let $z \in \tilde{S}^\perp$ have a positive component, $\lambda = \max_i z_i > 0$, and $\tilde{I} = \{i \mid z_i = \lambda\}$. Then z corresponds to a positive face $\tilde{f}$ of $\tilde{P}$ with $\tilde{I} = \{i \mid \tilde{w}^i \in \tilde{f}\}$. If the sign vector $\pi \in \{0,+\}^n$ with $\pi^+ = \tilde{I}$ satisfies $\pi \notin \operatorname{sign}(S)_\oplus$, then z is nondegenerate, by definition. $\square$

4 Robustness of bijectivity

We study the robustness of the simultaneous bijectivity of F_c for all $c > 0$ with respect to small perturbations of the exponents $\tilde{W}$ or/and the coefficients W, corresponding to small perturbations of the subspaces $\tilde{S}$ and S (in the Grassmannian).

The set of all $n-d$ dimensional subspaces S of $\mathbb{R}^n$ is the Grassmann manifold of rank $n-d$. It is a compact, connected smooth manifold of dimension $d(n-d)$, see e.g. [21, Chapter IV.7]. There are many metrics on the Grassmannian that generate the same topology, for example, two subspaces S and $\tilde{S}$ are close if and only if, for all $x \in S$ with $|x| = 1$, there exists $\tilde{x} \in \tilde{S}$ close to x, and the other way round.

4.1 Perturbations of the exponents

First, we consider small perturbations of the subspace $\tilde{S}$, corresponding to the exponents $\tilde{W}$ in F_c. As it turns out, the closure of $\operatorname{sign}(\tilde{S})$ plays an important role.

Definition 23. Let $T \subseteq \{-,0,+\}^n$. We define its *closure*

$$\overline{T} = \{\tau \in \{-,0,+\}^n \mid \tau \leq \rho \text{ for some } \rho \in T\}.$$

Clearly, $T_1 \subseteq \overline{T_2}$ implies $\overline{T_1} \subseteq \overline{T_2}$.

Lemma 24. *Let S be a subspace of $\mathbb{R}^n$ and S_ε be a small perturbation. Then $\operatorname{sign}(S) \subseteq \overline{\operatorname{sign}(S_\varepsilon)}$.*

Proof. Let $\pi \in \operatorname{sign}(S)$ and a corresponding $x \in S$ with $\pi = \operatorname{sign}(x)$. Then there is $x_\varepsilon \in S_\varepsilon$ close to x. For a small enough perturbation S_ε, nonzero components keep their signs (but zero components can become nonzero), that is, $\operatorname{sign}(x) \leq \operatorname{sign}(x_\varepsilon)$. Hence, $\pi \in \overline{\operatorname{sign}(S_\varepsilon)}$. $\square$

We start by studying injectivity.

Lemma 25. *Let $S, \tilde{S}$ be subspaces of $\mathbb{R}^n$. If $\operatorname{sign}(S) \cap \operatorname{sign}(\tilde{S}_\varepsilon^\perp) = \{0\}$ for all small perturbations $\tilde{S}_\varepsilon$, then $\operatorname{sign}(S) \subseteq \overline{\operatorname{sign}(\tilde{S})}$.*

Proof. Suppose $\operatorname{sign}(S) \subseteq \overline{\operatorname{sign}(\tilde{S})}$ does not hold. Then there is a nonzero sign vector $\pi \in \operatorname{sign}(S)$ with $\pi \notin \overline{\operatorname{sign}(\tilde{S})}$. We will find a small perturbation $\tilde{S}_\varepsilon$ such that $\pi \in \operatorname{sign}(\tilde{S}_\varepsilon^\perp)$ and hence $\operatorname{sign}(S) \cap \operatorname{sign}(\tilde{S}_\varepsilon^\perp) = \{0\}$ is violated.

By Corollary 52 in Appendix B, the nonexistence of $\rho \in \operatorname{sign}(\tilde{S})$ with $\rho \geq \pi$ implies the existence of a nonzero $\tilde{\tau} \in \operatorname{sign}(\tilde{S}^\perp)$ with $\tilde{\tau} \leq \pi$. If $\tilde{\tau} = \pi$, then $\pi \in \operatorname{sign}(\tilde{S}^\perp)$, as desired. Otherwise, let $\tilde{\tau} = \operatorname{sign}(x)$ for $x \in \tilde{S}^\perp$. We find a perturbation $x_\varepsilon = x + \varepsilon e$ with $\varepsilon > 0$ small and $e \in \mathbb{R}^n$ such that $\operatorname{sign}(x_\varepsilon) = \pi$. In particular, we choose $e_i = 1$ if $x_i = 0$ and $i \in \pi^+$, $e_i = -1$ if $x_i = 0$ and $i \in \pi^-$, and $e_i = 0$ otherwise. Then, we rescale x_ε such that $|x_\varepsilon| = |x|$. Finally, we find an orthogonal matrix $U \in \mathbb{R}^{n \times n}$ (close to the identity) such that $Ux = x_\varepsilon$. Then $x_\varepsilon = Ux \perp U\tilde{S} = \tilde{S}_\varepsilon$, that is, $x_\varepsilon \in \tilde{S}_\varepsilon^\perp$ and $\pi \in \operatorname{sign}(\tilde{S}_\varepsilon^\perp)$, as desired. ☐

Lemma 26. *Let $S, \tilde{S}$ be subspaces of $\mathbb{R}^n$. If $\operatorname{sign}(S) \subseteq \overline{\operatorname{sign}(\tilde{S})}$, then $\operatorname{sign}(S) \cap \operatorname{sign}(\tilde{S}^\perp) = \{0\}$.*

Proof. Assume there exists a nonzero $\tilde{\tau} \in \operatorname{sign}(S) \cap \operatorname{sign}(\tilde{S}^\perp)$. If $\operatorname{sign}(S) \subseteq \overline{\operatorname{sign}(\tilde{S})}$, then there exists $\rho \in \operatorname{sign}(\tilde{S})$ with $\rho \geq \tilde{\tau}$. In particular, $\tilde{\tau} \cdot \rho \neq 0$, thereby contradicting $\tilde{\tau} \in \operatorname{sign}(\tilde{S})^\perp = \operatorname{sign}(\tilde{S})^\perp$ and $\rho \in \operatorname{sign}(\tilde{S})$. Cf. Corollary 53 in Appendix B. ☐

Proposition 27. *Let $S, \tilde{S}$ be subspaces of $\mathbb{R}^n$. Then $\operatorname{sign}(S) \cap \operatorname{sign}(\tilde{S}_\varepsilon^\perp) = \{0\}$ for all small perturbations $\tilde{S}_\varepsilon$ if and only if $\operatorname{sign}(S) \subseteq \overline{\operatorname{sign}(\tilde{S})}$.*

Proof. ($\Rightarrow$): By Lemma 25.

($\Leftarrow$): Assume $\operatorname{sign}(S) \subseteq \overline{\operatorname{sign}(\tilde{S})}$. By Lemma 24, $\operatorname{sign}(\tilde{S}) \subseteq \overline{\operatorname{sign}(\tilde{S}_\varepsilon)}$ for all small perturbations $\tilde{S}_\varepsilon$ which implies $\overline{\operatorname{sign}(\tilde{S})} \subseteq \overline{\operatorname{sign}(\tilde{S}_\varepsilon)}$. Hence, $\operatorname{sign}(S) \subseteq \overline{\operatorname{sign}(\tilde{S}_\varepsilon)}$. By Lemma 26, $\operatorname{sign}(S) \cap \operatorname{sign}(\tilde{S}_\varepsilon^\perp) = \{0\}$. ☐

Corollary 28. *Let $S, \tilde{S}$ be subspaces of $\mathbb{R}^n$. Then*

$$\operatorname{sign}(S) \subseteq \overline{\operatorname{sign}(\tilde{S})} \quad \textit{if and only if} \quad \operatorname{sign}(S^\perp) \subseteq \overline{\operatorname{sign}(\tilde{S}^\perp)}.$$

Proof. By Corollary 5, $\operatorname{sign}(S) \cap \operatorname{sign}(\tilde{S}_\varepsilon^\perp) = \{0\}$ is equivalent to $\operatorname{sign}(S^\perp) \cap \operatorname{sign}(\tilde{S}_\varepsilon) = \{0\}$. By Proposition 27 twice, the former statement (for all small perturbations $\tilde{S}_\varepsilon$) is equivalent to $\operatorname{sign}(S) \subseteq \overline{\operatorname{sign}(\tilde{S})}$ and the latter to $\operatorname{sign}(S^\perp) \subseteq \overline{\operatorname{sign}(\tilde{S}^\perp)}$. ☐

In terms of the map F_c (and the associated subspaces S and $\tilde{S}$), Proposition 27 states that

$$\begin{array}{l} F_c \text{ is injective for all } c > 0 \\ \text{and all small perturbations } \tilde{S}_\varepsilon \end{array} \quad \Leftrightarrow \quad \operatorname{sign}(S) \subseteq \overline{\operatorname{sign}(\tilde{S})}.$$

In Proposition 29 and Theorem 31 below, we will show that

$$\operatorname{sign}(S) \subseteq \overline{\operatorname{sign}(\tilde{S})} \quad \Rightarrow \quad F_c \text{ is bijective for all } c > 0$$

and

$$\begin{array}{l} F_c \text{ is bijective for all } c > 0 \\ \text{and all small perturbations } \tilde{S}_\varepsilon \end{array} \quad \Leftrightarrow \quad \operatorname{sign}(S) \subseteq \overline{\operatorname{sign}(\tilde{S})}.$$

First, we prove that the closure condition

$$\operatorname{sign}(S) \subseteq \overline{\operatorname{sign}(\tilde{S})} \tag{cc}$$

implies the bijectivity of F_c for all $c > 0$, that is, conditions (i), (ii), and (iii) in Theorem 14. For an alternative proof, using differential topology, see [16].

Proposition 29. *If* $\operatorname{sign}(S) \subseteq \overline{\operatorname{sign}(\tilde{S})}$*, then the map* F_c *is a diffeomorphism for all* $c > 0$*.*

Proof. (cc) $\Rightarrow$ (i): By Lemma 26.

(cc) $\Rightarrow$ (ii):

Assume $\neg$(ii), that is, the existence of a nonzero $\tilde{\tau} \in \operatorname{sign}(\tilde{S}^\perp)_\oplus$ with $\tau \not\leq \tilde{\tau}$ for all nonzero $\tau \in \operatorname{sign}(S^\perp)_\oplus$, in fact, for all nonzero $\tau \in \operatorname{sign}(S^\perp)$. By Corollary 52 in Appendix B, the nonexistence of a nonzero $\tau \in \operatorname{sign}(S^\perp)$ with $\tau \leq \tilde{\tau}$ implies the existence of $\pi \in \operatorname{sign}(S)$ with $\pi \geq \tilde{\tau}$.

Now, if $\operatorname{sign}(S) \subseteq \overline{\operatorname{sign}(\tilde{S})}$, then there exists $\rho \in \operatorname{sign}(\tilde{S})$ with $\rho \geq \pi$ and hence $\rho \geq \tilde{\tau}$. In particular, $\tilde{\tau} \cdot \rho \neq 0$, thereby contradicting $\tilde{\tau} \in \operatorname{sign}(\tilde{S}^\perp) = \operatorname{sign}(\tilde{S})^\perp$ and $\rho \in \operatorname{sign}(\tilde{S})$.

(cc) $\Rightarrow$ (iv) in Proposition 21:

Assume $\neg$(iv), that is, the existence of $\tilde{\tau} \in \operatorname{sign}(\tilde{S}^\perp)$ with $\tilde{\tau}^+ \neq \emptyset$, $\pi \in \operatorname{sign}(S)_\oplus$ with $\pi^+ = \tilde{\tau}^+$, and $\rho \in \operatorname{sign}(S)$ with $\tilde{\tau}^+ \cup \tilde{\tau}^- \subseteq \rho^+$. By composition, $\pi' = \pi \circ (-\rho) \in \operatorname{sign}(S)$, where $\pi'_i = +$ for $i \in \tilde{\tau}^+$ and $\pi'_i = -$ for $i \in \tilde{\tau}^-$, that is, $\pi' \geq \tilde{\tau}$.

Now, if $\operatorname{sign}(S) \subseteq \overline{\operatorname{sign}(\tilde{S})}$, then there exists $\rho' \in \operatorname{sign}(\tilde{S})$ with $\rho' \geq \pi'$ and hence $\rho' \geq \tilde{\tau}$. In particular, $\tilde{\tau} \cdot \rho' \neq 0$, thereby contradicting $\tilde{\tau} \in \operatorname{sign}(\tilde{S}^\perp) = \operatorname{sign}(\tilde{S})^\perp$ and $\rho' \in \operatorname{sign}(\tilde{S})$. $\square$

However, the closure condition (cc) is not necessary for bijectivity. Recall that there is a (weakest) sign-vector condition sufficient for bijectivity, involving conditions (i), (ii), and (iv) in Proposition 21.

Example 30. Let F_c be given by the matrices

$$\tilde{W} = \begin{pmatrix} 1 & 0 & -1 \end{pmatrix} \quad \text{and} \quad W = \begin{pmatrix} 1 & 1 & -1 \end{pmatrix}.$$

Obviously, $\tilde{C} = C = \mathbb{R}$. Now, for $\tau = (+,+,-)^\mathsf{T} \in \operatorname{sign}(\operatorname{im} W^\mathsf{T}) = \operatorname{sign}(S^\perp)$, there is no $\tilde{\tau} \in \operatorname{sign}(\operatorname{im} \tilde{W}^\mathsf{T}) = \operatorname{sign}(\tilde{S}^\perp)$ with $\tilde{\tau} \geq \tau$. Hence, $\operatorname{sign}(S^\perp) \not\subseteq \overline{\operatorname{sign}(\tilde{S}^\perp)}$, that is, the closure condition (cc) does not hold. Still, there is no nonzero $\pi \in \operatorname{sign}(\ker W)_\oplus = \operatorname{sign}(S)_\oplus$, and hence condition (iv) holds. Further, conditions (i) and (ii) hold, and F_c is bijective for all $c > 0$.

In fact, the closure condition (cc) is equivalent to bijectivity for all small perturbations $\tilde{S}_\varepsilon$.

Theorem 31. *The map F_c is a diffeomorphism for all $c > 0$ and all small perturbations $\tilde{S}_\varepsilon$ if and only if* $\mathrm{sign}(S) \subseteq \overline{\mathrm{sign}(\tilde{S})}$.

Proof. By Lemma 24, $\mathrm{sign}(S) \subseteq \overline{\mathrm{sign}(\tilde{S})}$ implies $\mathrm{sign}(S) \subseteq \overline{\mathrm{sign}(\tilde{S}_\varepsilon)}$ for all small perturbations $\tilde{S}_\varepsilon$. By Proposition 29, the latter implies the bijectivity of F_c for all $c > 0$ and all small perturbations $\tilde{S}_\varepsilon$.

Bijectivity implies injectivity, that is, $\mathrm{sign}(S) \cap \mathrm{sign}(\tilde{S}_\varepsilon^\perp) = \{0\}$, for all small perturbations $\tilde{S}_\varepsilon$. By Lemma 25, the latter implies $\mathrm{sign}(S) \subseteq \overline{\mathrm{sign}(\tilde{S})}$. □

Corollary 4 relates *chirotopes* (signs of maximal minors of W and $\tilde{W}$) to *vectors* (sign vectors of $S = \ker W$ and $\tilde{S} = \ker \tilde{W}$). By varying over all small perturbations $\tilde{S}_\varepsilon$, we obtain the following result.

Proposition 32. *Let $S, \tilde{S}$ be subspaces of $\mathbb{R}^n$ of dimension $n - d$ (with $d \leq n$). For every $W, \tilde{W} \in \mathbb{R}^{d \times n}$ (with full rank d) such that $S = \ker W$ and $\tilde{S} = \ker \tilde{W}$, the following statements are equivalent.*

1. $\mathrm{sign}(S) \subseteq \overline{\mathrm{sign}(\tilde{S})}$.
2. $\det(W_I) \neq 0$ *implies* $\det(W_I) \det(\tilde{W}_I) > 0$ *for all subsets $I \subseteq [n]$ of cardinality d (or '< 0' for all I).*

Proof. By Proposition 27, statement 1 is equivalent to $\mathrm{sign}(S) \cap \mathrm{sign}(\tilde{S}_\varepsilon^\perp) = \{0\}$ for all small perturbations $\tilde{S}_\varepsilon$. By Corollary 4, this is equivalent to

> $\det(W_I) \det(\tilde{W}_{\varepsilon,I}) \geq 0$ for all $I \subseteq [n]$ of cardinality d (or '≤ 0' for all I)
> and $\det(W_I) \det(\tilde{W}_{\varepsilon,I}) \neq 0$ for some I,
> for all small perturbations $\tilde{W}_\varepsilon$ of $\tilde{W}$.

This is equivalent to statement 2, thereby using that $\det(\tilde{W}_I) = 0$ implies $\det(\tilde{W}_{\varepsilon_1,I}) < 0$ and $\det(\tilde{W}_{\varepsilon_2,I}) > 0$ for some small perturbations $\tilde{W}_{\varepsilon_1}$ and $\tilde{W}_{\varepsilon_2}$. □

Now we can extend Theorem 31. In particular, we can characterize the bijectivity of F_c for all $c > 0$ and all small perturbations $\tilde{S}_\varepsilon$ not only in terms of sign vectors, but also in terms of maximal minors.

Corollary 33. *The following statements are equivalent:*

1. *F_c is a diffeomorphism for all $c > 0$ and all small perturbations $\tilde{S}_\varepsilon$.*
2. $\mathrm{sign}(S) \subseteq \overline{\mathrm{sign}(\tilde{S})}$.
3. $\det(W_I) \neq 0$ *implies* $\det(W_I) \det(\tilde{W}_I) > 0$ *for all subsets $I \subseteq [n]$ of cardinality d (or '< 0' for all I).*

Proof. ($1 \Leftrightarrow 2$): By Theorem 31. ($2 \Leftrightarrow 3$): By Proposition 32. □

4.2 Perturbations of the coefficients

Next, we consider small perturbations of the subspace S, corresponding to the coefficients W in F_c. We start by studying injectivity. By Corollary 5, the perturbed injectivity condition $\mathrm{sign}(S_\varepsilon) \cap \mathrm{sign}(\tilde{S}^\perp) = \{0\}$ is equivalent to $\mathrm{sign}(\tilde{S}) \cap \mathrm{sign}(S_\varepsilon^\perp) = \{0\}$. By exchanging the roles of S and $\tilde{S}$ in Proposition 27, we immediately obtain the desired result.

Corollary 34. *Let $S, \tilde{S}$ be subspaces of $\mathbb{R}^n$. Then $\mathrm{sign}(S_\varepsilon) \cap \mathrm{sign}(\tilde{S}^\perp) = \{0\}$ for all small perturbations S_ε if and only if $\mathrm{sign}(\tilde{S}) \subseteq \overline{\mathrm{sign}(S)}$.*

The closure condition

$$\mathrm{sign}(\tilde{S}) \subseteq \overline{\mathrm{sign}(S)} \tag{cc'}$$

is equivalent to $\mathrm{sign}(\tilde{S}^\perp) \subseteq \overline{\mathrm{sign}(S^\perp)}$, by Corollary 28. As opposed to (cc), it does not imply bijectivity, in fact, it implies conditions (i) and (iii) in Theorem 14, but not condition (ii).

Proposition 35. *If $\mathrm{sign}(\tilde{S}) \subseteq \overline{\mathrm{sign}(S)}$, then conditions (i) and (iii) in Theorem 14 hold.*

Proof. (cc') $\Rightarrow$ (i): By Corollary 34.

(cc') $\Rightarrow$ (iv) in Proposition 21:

Assume $\neg$(iv) and hence the existence of $\tilde{\tau} \in \mathrm{sign}(\tilde{S}^\perp)$ and $\pi \in \mathrm{sign}(S)_\oplus$ with $\tilde{\tau}^+ = \pi^+ \neq \emptyset$, in particular, $\tilde{\tau} \geq \pi$. Now, if $\mathrm{sign}(\tilde{S}) \subseteq \overline{\mathrm{sign}(S)}$, then there exists $\rho \in \mathrm{sign}(S^\perp)$ with $\rho \geq \tilde{\tau}$ and hence $\rho \geq \pi$. In particular, $\pi \cdot \rho \neq 0$, thereby contradicting $\pi \in \mathrm{sign}(S)$ and $\rho \in \mathrm{sign}(S^\perp) = \mathrm{sign}(S)^\perp$. $\square$

Example 36. Let F_c be given by the matrices

$$\tilde{W} = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \quad \text{and} \quad W = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}.$$

Obviously, $\tilde{C} = \mathbb{R} \times \mathbb{R}_{\geq 0}$ and $C = \mathbb{R}^2_{\geq 0}$. Now, $\tilde{S} = \ker \tilde{W} = \mathrm{im}(1,0,1)^\mathsf{T}$, $S = \ker W = \mathrm{im}(1,-1,1)^\mathsf{T}$, and hence $\mathrm{sign}(\tilde{S}) \subseteq \overline{\mathrm{sign}(S)}$. However, $\mathrm{sign}(\tilde{S}^\perp)_\oplus = \{(0,0,0)^\mathsf{T},\ (0,+,0)^\mathsf{T}\}$, $\mathrm{sign}(S^\perp)_\oplus = \{(0,0,0)^\mathsf{T},\ (0,+,+)^\mathsf{T},\ (+,+,0)^\mathsf{T}\}$, and hence condition (ii) does not hold.

Interestingly, conditions (cc') and (ii) imply the equality of the face lattices of C and $\tilde{C}$.

Proposition 37. *If $\mathrm{sign}(\tilde{S}) \subseteq \overline{\mathrm{sign}(S)}$ and condition (ii) in Theorem 14 holds, then $\mathrm{sign}(S^\perp)_\oplus = \mathrm{sign}(\tilde{S}^\perp)_\oplus$.*

Proof. Recall that, by the proof of Proposition 29, (cc) implies (ii); analogously, (cc') implies

(ii') for every nonzero $\tau \in \mathrm{sign}(S^\perp)_\oplus$, there is a nonzero $\tilde{\tau} \in \mathrm{sign}(\tilde{S}^\perp)_\oplus$ such that $\tilde{\tau} \leq \tau$.

On the one hand, let $\tau \in \mathrm{sign}(S^\perp)_\oplus$ have minimal support. By (ii'), there is a nonzero $\tilde\tau \in \mathrm{sign}(\tilde S^\perp)_\oplus$ such that $\tilde\tau \le \tau$. By (ii), there is a nonzero $\tau' \in \mathrm{sign}(S^\perp)_\oplus$ such that $\tau' \le \tilde\tau$. Altogether, $\tau' \le \tilde\tau \le \tau$. Now, $\tau' = \tau$, since τ has minimal support, and hence $\tilde\tau = \tau$. That is, there is a *unique* nonzero $\tilde\tau \in \mathrm{sign}(\tilde S^\perp)_\oplus$ (namely $\tilde\tau = \tau$) such that $\tilde\tau \le \tau$. In particular, $\tilde\tau$ has minimal support.

On the other hand, let $\tilde\tau \in \mathrm{sign}(\tilde S^\perp)_\oplus$ have minimal support. By an analogous argument, there is a *unique* nonzero $\tau \in \mathrm{sign}(S^\perp)_\oplus$ (namely $\tau = \tilde\tau$) such that $\tau \le \tilde\tau$. In particular, $\tilde\tau$ has minimal support. Hence, elements of $\mathrm{sign}(S^\perp)_\oplus$ and $\mathrm{sign}(\tilde S^\perp)_\oplus$ with minimal support are in one-to-one correspondence. Finally, every nonzero, nonnegative sign vector of a subspace is the composition of nonnegative sign vectors with minimal support, cf. Theorem 49 in Appendix A. Hence, $\mathrm{sign}(S^\perp)_\oplus = \mathrm{sign}(\tilde S^\perp)_\oplus$. □

It remains to study the robustness of condition (ii).

Lemma 38. *If, for all small perturbations S_ε, the map F_c is surjective and condition (ii) in Theorem 14 holds, then either $C = \tilde C = \mathbb{R}^d$ or $(+,\dots,+)^\mathsf{T} \in \mathrm{sign}(S^\perp) \cap \mathrm{sign}(\tilde S^\perp)$.*

Proof. If neither $C = \mathbb{R}^d$ nor $(+,\dots,+)^\mathsf{T} \in \mathrm{sign}(S^\perp)$, then C has a nontrivial lineality space. On the one hand, there is a small perturbation S_{ε_1} such that[1] $C_{\varepsilon_1} = \mathbb{R}^d$; hence $\tilde C = \mathbb{R}^d$, by (ii). On the other hand, there is a small perturbation S_{ε_2} such that $(+,\dots,+)^\mathsf{T} \in \mathrm{sign}(S_{\varepsilon_2}^\perp)$; hence $(+,\dots,+)^\mathsf{T} \in \mathrm{sign}(\tilde S^\perp)$, by Proposition 19. A contradiction.

If $C = \mathbb{R}^d$, then $\tilde C = \mathbb{R}^d$, by (ii). If $(+,\dots,+)^\mathsf{T} \in \mathrm{sign}(S^\perp)$, then $(+,\dots,+)^\mathsf{T} \in \mathrm{sign}(\tilde S^\perp)$, by Proposition 19. □

That is, condition (ii) is robust only in two extreme cases regarding the geometry of $C = \mathrm{cone}(W)$. We consider the case $(+,\dots,+)^\mathsf{T} \in \mathrm{sign}(S^\perp)$ separately.

We call C *robustly generated* if either $d = 1$ or, on every extreme ray of C, there lies a unique vector w^i, and all other vectors lie in the interior. In terms of sign vectors, C is robustly generated if

> a nonzero $\tau \in \mathrm{sign}(S^\perp)_\oplus$ has minimal support if and only if, for every $i \in \tau^0$, there exists $\hat\tau \in \mathrm{sign}(S^\perp)_\oplus$ with $\hat\tau^0 = \{i\}$.

In this case, $\mathrm{sign}(S_\varepsilon^\perp)_\oplus = \mathrm{sign}(S^\perp)_\oplus$ for all small perturbations S_ε, and condition (ii) is robust. In fact, (ii) being robust implies C being robustly generated.

Lemma 39. *Let $(+,\dots,+)^\mathsf{T} \in \mathrm{sign}(S^\perp)$ and $\mathrm{sign}(S^\perp)_\oplus = \mathrm{sign}(\tilde S^\perp)_\oplus$. If condition (ii) in Theorem 14 holds for all small perturbations S_ε, then C and $\tilde C$ are robustly generated.*

---

[1] Let $L \subset [n]$ be the indices of the vectors w^i in the lineality space and $I = [n] \setminus L$. Hence, there are $c_i > 0$ for $i \in L$ such that $\sum_{i\in L} c_i w^i = 0$ and $\sum_{i\in L} c_i = 1$. Consider small perturbations S_ε as follows: $w^i_\varepsilon = w^i$ for $i \in I$ and $w^i_\varepsilon = w^i - \varepsilon \sum_{j\in I} w^j$ for $i \in L$, where $\varepsilon > 0$. Then, $\sum_{i\in L} c_i w^i_\varepsilon + \sum_{i \in I} \varepsilon\, w^i_\varepsilon = 0$, and hence $(+,\dots,+)^\mathsf{T} \in \mathrm{sign}(\ker W_\varepsilon) = \mathrm{sign}(S_\varepsilon)$, that is, $C_\varepsilon = \mathbb{R}^d$.

Proof. Let $d > 1$. Assume that C is not robustly generated, and let f be a maximal proper face, characterized by $\tau \in \operatorname{sign}(S^\perp)_\oplus$ with minimal support, such that $w^j \in f$ for some $j \in [n]$, but w^j is not needed to generate f. Further let $\tilde{f}$ be the corresponding maximal proper face of $\tilde{C}$, characterized by $\tilde{\tau} = \tau \in \operatorname{sign}(\tilde{S}^\perp)_\oplus$ with minimal support. In particular, $\tau_j = \tilde{\tau}_j = 0$.

Now, consider a small perturbation S_ε such that $w_\varepsilon^j \in C^\circ$ and $w_\varepsilon^i = w^i$ for $i \neq j$ (and hence $C_\varepsilon = C$). Then, $\tau'_j = +$ for all $\tau' \in \operatorname{sign}(S_\varepsilon^\perp)_\oplus$, and there is no $\tau' \in \operatorname{sign}(S_\varepsilon^\perp)_\oplus$ with $\tau' \leq \tilde{\tau}$, contradicting (ii) for $\tilde{\tau}$. □

Finally, the closure condition (cc') together with sign-vector conditions regarding the geometry of the cones C and $\tilde{C}$ is equivalent to bijectivity for all small perturbations S_ε.

Theorem 40. *The map F_c is a diffeomorphism for all $c > 0$ and all small perturbations S_ε if and only if* $\operatorname{sign}(\tilde{S}) \subseteq \overline{\operatorname{sign}(S)}$ *and*

either $C = \tilde{C} = \mathbb{R}^d$
or $(+, \ldots, +)^\mathsf{T} \in \operatorname{sign}(S^\perp) \cap \operatorname{sign}(\tilde{S}^\perp)$, $\operatorname{sign}(S^\perp)_\oplus = \operatorname{sign}(\tilde{S}^\perp)_\oplus$, and C and $\tilde{C}$ are robustly generated.

Proof. By Theorem 14, the simultaneous bijectivity of F_c for all $c > 0$ is equivalent to conditions (i), (ii), and (iii) in Theorem 14.

By Corollary 34, condition (i), that is, $\operatorname{sign}(S_\varepsilon) \cap \operatorname{sign}(\tilde{S}^\perp) = \{0\}$, for all small perturbations S_ε, is equivalent to $\operatorname{sign}(\tilde{S}) \subseteq \overline{\operatorname{sign}(S)}$.

Now assume conditions (i), (ii), and (iii), for all small perturbations S_ε. By Proposition 37, $\operatorname{sign}(S^\perp)_\oplus = \operatorname{sign}(\tilde{S}^\perp)_\oplus$. By Lemma 38, either $C = \tilde{C} = \mathbb{R}^d$ or $(+, \ldots, +)^\mathsf{T} \in \operatorname{sign}(S^\perp) \cap \operatorname{sign}(\tilde{S}^\perp)$. In the latter case, by Lemma 39, C and $\tilde{C}$ are robustly generated.

Conversely, $\tilde{C} = \mathbb{R}^d$ (that is, $\operatorname{sign}(\tilde{S}^\perp)_\oplus = \{0\}$) trivially implies condition (ii) for all small perturbations S_ε. By Lemma 24, $\operatorname{sign}(\tilde{S}) \subseteq \overline{\operatorname{sign}(S)}$ implies $\operatorname{sign}(\tilde{S}) \subseteq \overline{\operatorname{sign}(S_\varepsilon)}$ for all small perturbations S_ε, and by Proposition 35 (for $\tilde{S}$ and S_ε), this implies condition (iii) for all small perturbations S_ε.

Finally, $(+, \ldots, +)^\mathsf{T} \in \operatorname{sign}(S^\perp)$, $\operatorname{sign}(S^\perp)_\oplus = \operatorname{sign}(\tilde{S}^\perp)_\oplus$, and C being robustly generated imply $(+, \ldots, +)^\mathsf{T} \in \operatorname{sign}(S_\varepsilon^\perp)$ and hence condition (iii), for all small perturbations S_ε. Further, they imply $\operatorname{sign}(S_\varepsilon^\perp)_\oplus = \operatorname{sign}(\tilde{S}^\perp)_\oplus$ and hence condition (ii), for all small perturbations S_ε. □

4.3 General perturbations

Finally, we consider small perturbations of both subspaces, S and $\tilde{S}$, corresponding to the coefficients W and the exponents $\tilde{W}$ in F_c.

The next result relates *chirotopes* to *cocircuits* (sign vectors of $S^\perp = \operatorname{im} W^\mathsf{T}$ and $\tilde{S}^\perp = \operatorname{im} \tilde{W}^\mathsf{T}$ with minimal support).

Lemma 41. *Let $S, \tilde{S}$ be subspaces of $\mathbb{R}^n$ of dimension $n - d$ (with $d \leq n$). For every $W, \tilde{W} \in \mathbb{R}^{d \times n}$ (with full rank d) such that $S = \ker W$ and $\tilde{S} = \ker \tilde{W}$, the following statements are equivalent.*

1. $\operatorname{sign}(S) = \operatorname{sign}(\tilde{S})$, *and a nonzero* $\tau \in \operatorname{sign}(S^\perp)$ *has minimal support if and only if* $|\tau^0| = d - 1$.

2. $\det(W_I)\det(\tilde{W}_I) > 0$ *for all subsets* $I \subseteq [n]$ *of cardinality* d *(or '*< 0*' for all* I*).*

Proof. By using the standard chirotope/cocircuit translation for subspaces of $\mathbb{R}^n$, see Theorem 48 in Appendix A. ☐

As it turns out, all maximal minors of W and $\tilde{W}$ being nonzero and having matching signs is equivalent to bijectivity for all small perturbations S_ε and $\tilde{S}_{\tilde{\varepsilon}}$.

Theorem 42. *The following statements are equivalent:*

1. F_c *is a diffeomorphism for all* $c > 0$ *and all small perturbations* S_ε *and* $\tilde{S}_{\tilde{\varepsilon}}$.

2. $\operatorname{sign}(S) = \operatorname{sign}(\tilde{S})$, *and a nonzero* $\tau \in \operatorname{sign}(S^\perp)$ *has minimal support if and only if* $|\tau^0| = d - 1$.

3. $\det(W_I)\det(\tilde{W}_I) > 0$ *for all subsets* $I \subseteq [n]$ *of cardinality* d *(or '*< 0*' for all* I*).*

Proof. ($1 \Rightarrow 3$): Statement 1 implies the injectivity of F_c for all $c > 0$, that is, $\operatorname{sign}(S_\varepsilon) \cap \operatorname{sign}(\tilde{S}_{\tilde{\varepsilon}}^\perp) = \{0\}$, for all small perturbations S_ε, $\tilde{S}_{\tilde{\varepsilon}}$. By Corollary 4, this is equivalent to

> $\det(W_{\varepsilon,I})\det(\tilde{W}_{\tilde{\varepsilon},I}) \geq 0$ for all $I \subseteq [n]$ of cardinality d (or '≤ 0' for all I)
> and $\det(W_{\varepsilon,I})\det(\tilde{W}_{\tilde{\varepsilon},I}) \neq 0$ for some I,
> for all small perturbations W_ε of W and $\tilde{W}_{\tilde{\varepsilon}}$ of $\tilde{W}$.

This is equivalent to statement 3.

($3 \Rightarrow 1$): Statement 3 implies

> $\det(W_{\varepsilon,I})\det(\tilde{W}_{\tilde{\varepsilon},I}) > 0$ for all $I \subseteq [n]$ of cardinality d (or '< 0' for all I),
> for all small perturbations W_ε, $\tilde{W}_{\tilde{\varepsilon}}$.

By Lemma 41, this implies $\operatorname{sign}(S_\varepsilon) = \operatorname{sign}(\tilde{S}_{\tilde{\varepsilon}})$ and hence $\operatorname{sign}(S_\varepsilon) \subseteq \overline{\operatorname{sign}(\tilde{S}_{\tilde{\varepsilon}})}$, for all small perturbations W_ε, $\tilde{W}_{\tilde{\varepsilon}}$. By Proposition 29, this implies statement 1.

($2 \Leftrightarrow 3$): By Lemma 41. ☐

By Theorem 40, bijectivity for all $c > 0$ and all small perturbations S_ε already implies that either $C = \tilde{C} = \mathbb{R}^d$ or $(+, \ldots, +)^\mathsf{T} \in \operatorname{sign}(S^\perp) \cap \operatorname{sign}(\tilde{S}^\perp)$. In Theorem 42, this follows from the second part of condition 2. Assume that C has a nontrivial lineality space of dimension ℓ, generated by at least $\ell + 1$ vectors w^i. Then, a maximal proper face, having dimension $d - 1 = \ell + d'$, is generated by at least $(\ell + 1) + d' = d$ vectors and corresponds to a sign vector $\tau \in \operatorname{sign}(S^\perp)$ with minimal support, but $|\tau^0| \geq d$.

5 Applications to Chemical Reaction Networks

As mentioned in the introduction, our work is motivated by the study of chemical reaction networks with generalized mass-action kinetics. We present a derivation of our main problem (the characterization of bijectivity of families of exponential maps) and applications of our main results, in particular, Theorems 14 and 31.

We start with an introduction to chemical reaction networks (with mass-action kinetics). Thereby, we follow the graph-based approach introduced in [39]; see also [16, 33].

Consider the chemical reaction $1\mathsf{A} + 1\mathsf{B} \to \mathsf{C}$ (with *stoichiometric coefficients* equal to 1). Under the assumption of mass-action kinetics (MAK), the reaction rate is given by $v = k\, x_\mathsf{A}^1 x_\mathsf{B}^1$ (with *kinetic orders* equal to 1), where $k > 0$ is the *rate constant* and $x_\mathsf{A}, x_\mathsf{B} \geq 0$ are the concentrations of the chemical species A, B. Most importantly, the stoichiometric coefficients determine the kinetic orders. Given n species, a general reaction is written as $y \to y'$, where $y, y' \in \mathbb{R}^n_{\geq 0}$ are called (educt and product) *complexes*, and its rate is given by $v = k\, x^y$, where $x^y = \prod_{i=1}^n {x_i}^{y_i}$ is a monomial in the species concentrations $x \in \mathbb{R}^n_{\geq 0}$. In a network, an individual reaction $y \to y'$ contributes to the ODE for the species concentrations as $\frac{\mathrm{d}x}{\mathrm{d}t} = k\, x^y (y' - y) + \ldots$. Let $x = (x_\mathsf{A}, x_\mathsf{B}, x_\mathsf{C}, x_\mathsf{D}, \ldots)^\mathsf{T}$. For the reaction $\mathsf{A} + \mathsf{B} \to \mathsf{C}$ above, one has $y = (1, 1, 0, 0, \ldots)^\mathsf{T}$, $y' = (0, 0, 1, 0, \ldots)^\mathsf{T}$ and hence $x^y = x_\mathsf{A} x_\mathsf{B}$, $y' - y = (-1, -1, 1, 0, \ldots)^\mathsf{T}$.

A *chemical reaction network* (CRN) is based on a directed graph $G = (V, E)$. Every vertex $i \in V = \{1, \ldots, m\}$ is labeled with a complex $y(i) \in \mathbb{R}^n_{\geq 0}$, and every edge $i \to i' \in E$ (representing a reaction) is labeled with a rate constant $k_{i \to i'} > 0$. From the labeled digraph, one obtains the ODE for the species concentrations,

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \sum_{i \to i' \in E} k_{i \to i'}\, x^{y(i)} \big(y(i') - y(i)\big).$$

The sum ranges over all reactions, and every summand is a product of the reaction rate and the difference of product and educt complexes. The right-hand-side can be decomposed into stoichiometric and graphical contributions,

$$\frac{\mathrm{d}x}{\mathrm{d}t} = Y I_E\, v_k(x) = Y A_k\, x^Y,$$

where $Y \in \mathbb{R}^{n \times V}_{\geq 0}$ is the matrix of complexes, $I_E \in \mathbb{R}^{V \times E}$ is the incidence matrix, and $A_k \in \mathbb{R}^{V \times V}$ is the Laplacian matrix of the digraph G, labeled with the rate constants $k \in \mathbb{R}^E_{>0}$. The vector of reaction rates $v_k(x) \in \mathbb{R}^E_{\geq}$ is defined via $(v_k(x))_{i \to i'} = k_{i \to i'}\, x^{y(i)}$, and the vector of monomials $x^Y \in \mathbb{R}^V_{\geq 0}$ is defined via $(x^Y)_i = x^{y(i)}$, where $y(i)$ is the i-th column of Y.

A positive steady state $x \in \mathbb{R}^n_{>0}$ of the ODE that fulfills

$$A_k\, x^Y = 0$$

is called a *complex-balanced equilibrium*. Another important object is the *stoichiometric subspace*

$$S = \mathrm{im}(Y I_E).$$

Clearly, $\frac{\mathrm{d}x}{\mathrm{d}t} \in S$, and hence $x(t) \in x(0) + S$. For $x' \in \mathbb{R}^n_{\geq 0}$, the set $(x' + S) \cap \mathbb{R}^n_{\geq 0}$ is called a *stoichiometric class*. The *deficiency* of a CRN is given by

$$\delta = \dim(\ker Y \cap \operatorname{im} I_E) = m - \ell - \dim(S),$$

where m is the number of vertices, and ℓ is the number of connected components of the digraph. Finally, a CRN is called *weakly reversible* if all components of the digraph are strongly connected.

Now, we can state the celebrated deficiency zero theorem for MAK, formulated by Horn, Jackson, and Feinberg in 1972.

Theorem 43 ($\delta = 0$ theorem; cf. [28], [27], and [18])**.** *For a CRN with MAK, there exists a unique (complex-balanced, asymptotically stable) equilibrium in every stoichiometric class and for all rate constants if and only if $\delta = 0$ and the network is weakly reversible.*

The $\delta = 0$ theorem is a strong result. It characterizes CRNs with MAK that are dynamically as simple and stable as possible. However, MAK is an assumption that holds for elementary reactions in homogeneous and dilute solutions. In intracellular environments, which are highly structured and crowded, and for reaction mechanisms, more general kinetics are needed. As a prominent approach, biochemical systems theory [45, 51] proposes power laws in the species concentrations, where the kinetic orders may differ from the stoichiometric coefficients. In chemical reaction network theory, power-law kinetics has been termed general(ized) mass-action kinetics (GMAK) [28, 38, 39]. As already noted by Horn and Jackson [28], every CRN with GMAK can be written as another CRN with MAK, where the stoichiometric coefficients need not be integers. However, the resulting network typically loses desired properties such as weak reversibility and zero deficiency. In our more recent definition of CRNs with GMAK [38, 39], we allow for power-law kinetics, without having to rewrite the network.

In fact, a CRN with MAK may not have zero deficiency and may not be weakly reversible, but there may be a dynamically equivalent CRN with GMAK that has the desired properties. In particular, dynamical equivalence to a network having zero 'effective' and 'kinetic' deficiencies allows a parametrization of all positive equilibria [33]. Such a parametrization can be computed by linear algebra techniques and does not require tools from algebraic geometry such as Gröbner bases, as demonstrated for the EnvZ-OmpR and shuttled WNT signaling pathways. For algorithmic methods to identify dynamically equivalent CRNs and further applications to biochemical networks, see [30, 31, 50, 32].

Relations between biochemical systems theory and chemical reaction network theory are discussed in [3, 2, 49]. Power-law systems from biochemical systems theory can be realized as CRNs with GMAK having desired properties, and results e.g. from [38, 39] are applied to models of yeast fermentation, purine metabolism [3], and further paradigmatic models from systems biology [2].

We continue our introduction to chemical reaction networks (with generalized mass-action kinetics). For the reaction above, $1\mathsf{A} + 1\mathsf{B} \to \mathsf{C}$, now under the assumption of GMAK, the reaction rate is given by $v = k\, x_A^a x_B^b$, where the kinetic orders $a, b \in \mathbb{R}$ need not coincide with the stoichiometric coefficients.

One writes

$$\boxed{\begin{array}{c}1\mathsf{A}+1\mathsf{B}\\(a\mathsf{A}+b\mathsf{B})\end{array}}\to\boxed{\begin{array}{c}\mathsf{C}\\(\ldots)\end{array}}$$

with the kinetic-order information in brackets. For a general reaction

$$\boxed{\begin{array}{c}y\\(\tilde{y})\end{array}}\to\boxed{\begin{array}{c}y'\\(\ldots)\end{array}},$$

one has

$$v = k\, x^{\tilde{y}},$$

where $\tilde{y} \in \mathbb{R}^n$ is called a *kinetic(-order) complex.*

As above, a CRN is based on a digraph $G = (V, E)$, but now every vertex $i \in V$ is labeled with stoichiometric *and* kinetic-order complexes, $y(i)$ and $\tilde{y}(i)$, respectively. (And every edge is labeled with a rate constant.) From the labeled digraph, one obtains the ODE

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \sum_{i \to i' \in E} k_{i \to i'}\, x^{\tilde{y}(i)} \big(y(j) - y(i)\big).$$

Again the right-hand-side of the ODE can be decomposed, now into stoichiometric, graphical, and kinetic-order contributions,

$$\frac{\mathrm{d}x}{\mathrm{d}t} = Y A_k\, x^{\tilde{Y}},$$

where $\tilde{Y} \in \mathbb{R}^{n \times V}_{\geq 0}$ is the matrix of kinetic-order complexes. Accordingly, a steady state $x \in \mathbb{R}^n_{>0}$ that fulfills

$$A_k\, x^{\tilde{Y}} = 0$$

is called a complex-balanced equilibrium. Finally, like the corresponding stoichiometric objects, one introduces the *kinetic-order subspace*

$$\tilde{S} = \operatorname{im}(\tilde{Y} I_E)$$

and the *kinetic(-order) deficiency*

$$\tilde{\delta} = \dim(\ker \tilde{Y} \cap \operatorname{im} I_E) = m - \ell - \dim(\tilde{S}).$$

The classical $\delta = 0$ theorem holds for MAK. In previous work, we formulated a first analogue for GMAK.

Theorem 44 ($\tilde{\delta} = 0$ theorem; cf. [39])**.** *For a CRN with GMAK, there exists a complex-balanced equilibrium for all rate constants if and only if $\tilde{\delta} = 0$ and the network is weakly reversible.*

However, this theorem does not fully correspond to the classical one which guarantees the unique existence of a complex-balanced equilibrium in every stoichiometric class. For GMAK, complex-balanced equilibria are determined by kinetic orders, whereas classes are determined by stoichiometry. In fact, a true analogue requires extra conditions on the stoichiometric and kinetic-order subspaces, S and $\tilde{S}$.

For given $k \in \mathbb{R}^E_{>0}$, let Z_k be the set of complex-balanced equilibria, and for given $x' \in \mathbb{R}^n_{>0}$, let $(x'+S) \cap \mathbb{R}^n_{\geq 0}$ be the corresponding stoichiometric class. We aim to characterize existence and uniqueness of an element in the intersection

$$Z_k \cap (x'+S)$$

for all $x' \in \mathbb{R}^n_{>0}$, for all $k \in \mathbb{R}^E_{>0}$. By Theorem 44, $Z_k \neq \emptyset$ for all $k \in \mathbb{R}^E_{>0}$ if and only if $\tilde{\delta} = 0$ and the network is weakly reversible, which we assume in the following.

By Theorem 1 in [39], $x^*_k \in Z_k$ implies the exponential parametrization

$$Z_k = x^*_k \circ \mathrm{e}^{\tilde{S}^\perp} .$$

Moreover, for a weakly reversible CRN, every $x^* \in \mathbb{R}^n_{>0}$ is a complex-balanced equilibrium for some rate constant $k \in \mathbb{R}^E_{>0}$, see e.g. the proof of Lemma 1 in [39]. Hence, we aim to characterize existence and uniqueness of an element in the intersection

$$x^* \circ \mathrm{e}^{\tilde{S}^\perp} \cap (x'+S)$$

for all $x', x^* \in \mathbb{R}^n_{>0}$.

For fixed x', x^*, we are interested in existence and uniqueness of $u \in S$, $v \in \tilde{S}^\perp$ such that

$$x^* \circ \mathrm{e}^v = x' + u$$

and introduce $W \in \mathbb{R}^{d \times n}, \tilde{W} \in \mathbb{R}^{\tilde{d} \times n}$ with full ranks $d, \tilde{d} \leq n$ such that

$$S = \ker W, \quad \tilde{S} = \ker \tilde{W}.$$

We multiply with W, write $v = \tilde{W}^\mathsf{T} \xi$ with $\xi \in \mathbb{R}^{\tilde{d}}$, and obtain

$$W(x^* \circ \mathrm{e}^{\tilde{W}^\mathsf{T} \xi}) = W x'.$$

Hence, we are interested in existence and uniqueness of $\xi \in \mathbb{R}^{\tilde{d}}$ such that the last equation holds.

Finally, we note that $Wx' \in C^\circ$, the interior of $C = \operatorname{cone} W$, and vary over all $x' \in \mathbb{R}^n_{>0}$ or, equivalently, over all elements of C°. As a result, we aim to characterize bijectivity of the map

$$\begin{aligned} F_{x^*} \colon \mathbb{R}^{\tilde{d}} &\to C^\circ \subseteq \mathbb{R}^d, \\ \xi &\mapsto W(x^* \circ \mathrm{e}^{\tilde{W}^\mathsf{T} \xi}) = \sum_{i=1}^n x^*_i \, \mathrm{e}^{\tilde{w}^i \cdot \xi} \, w^i \end{aligned}$$

for all $x^* \in \mathbb{R}^n_{>0}$, that is, the simultaneous bijectivity of the map F_{x^*} for all $x^* > 0$. Indeed, this is the content of Theorem 14, and the deficiency zero theorem can be fully extended to GMAK (except for stability).

Theorem 45 ($\delta = \tilde{\delta} = 0$ theorem)**.** *For a CRN with GMAK, there exists a unique complex-balanced equilibrium in every stoichiometric class and for all rate constants if and only if $\delta = \tilde{\delta} = 0$, the network is weakly reversible, and conditions (i), (ii), (iii) in Theorem 14 hold.*

In contrast to MAK, where complex-balanced equilibria are asymptotically stable, already two-species CRNs with GMAK lead to planar systems which have a unique (complex-balanced) equilibrium, but show rich dynamical behavior, including super/sub-critical or degenerate Hopf bifurcations, centers, and up to three limit cycles, see [9, 10, 11, 8].

By Theorem 31 (and the problem derivation given above), Theorem 45 is robust with respect to small perturbations of the kinetic orders if and only if the closure condition $\operatorname{sign}(S) \subseteq \overline{\operatorname{sign}(\tilde{S})}$ holds.

Theorem 46 (robust $\delta = \tilde{\delta} = 0$ theorem)**.** *For a CRN with GMAK, there exists a unique complex-balanced equilibrium in every stoichiometric class, for all rate constants, and for all small perturbations of the kinetic orders if and only if $\delta = \tilde{\delta} = 0$, the network is weakly reversible, and $\operatorname{sign}(S) \subseteq \overline{\operatorname{sign}(\tilde{S})}$.*

For a CRN with MAK, the stoichiometric and kinetic-order subspaces agree, that is, $S = \tilde{S}$, and obviously $\operatorname{sign}(S) \subseteq \overline{\operatorname{sign}(\tilde{S})}$. Hence, the classical deficiency zero theorem for MAK is robust with respect to small perturbations of the kinetic orders (from the stoichiometric coefficients).

Corollary 47 (robust $\delta = 0$ theorem)**.** *For a CRN with MAK, there exists a unique (complex-balanced, asymptotically stable) equilibrium in every stoichiometric class, for all rate constants, and for all small perturbations of the kinetic orders (from the stoichiometric coefficients) if and only if $\delta = 0$ and the network is weakly reversible.*

Acknowledgments

The closure condition (cc) was suggested by Gheorghe Craciun as a criterion for the bijectivity of the family of exponential maps. We thank Gheorghe Craciun, Casian Pantea, and Polly Yu for fruitful discussions (at workshops at the University of Wisconsin-Madison in 2015, the American Institute of Mathematics, San Jose, in 2016, and the Banff International Research Station and the Mathematisches Forschungsinstitut Oberwolfach in 2017). We also thank three anonymous referees for their careful reading and numerous helpful comments.

SM was supported by the Austrian Science Fund (FWF), project P28406. GR was supported by the FWF, project P27229.

Appendices

A Sign vectors and face lattices

In the context of (realizable) oriented matroids, we discuss the relation between sign vectors of linear subspaces and face lattices of polyhedral cones. For further details, we refer to [4, Chapter 7], [52, Chapters 2 and 6], [42], and the encyclopedic study [7].

Let $W = (w^1, \ldots, w^n) \in \mathbb{R}^{d \times n}$ with $d \leq n$ have full rank. Then W is called a *vector configuration* (of n vectors in $\mathbb{R}^d$), and $\operatorname{im} W^\mathsf{T} \subseteq \mathbb{R}^n$ is a corresponding linear subspace. Now let $v = W^\mathsf{T} x \in \operatorname{im} W^\mathsf{T}$ with $x \in \mathbb{R}^d$. Then $v_i = w^i \cdot x$, and the sign vector $\tau = \operatorname{sign}(v) \in \operatorname{sign}(\operatorname{im} W^\mathsf{T}) \subseteq \{-, 0, +\}^n$ describes the positions of the vectors $w^1, \ldots, w^n$ relative to the hyperplane with normal vector x.

Elements of $\operatorname{sign}(\operatorname{im} W^\mathsf{T})$ are called *covectors*, and elements of $\operatorname{sign}(\operatorname{im} W^\mathsf{T})$ with minimal support are called *cocircuits*. Analogously, elements of $\operatorname{sign}(\ker W)$ are called *vectors*, and elements of $\operatorname{sign}(\ker W)$ with minimal support are called *circuits*.

The *chirotope* of the vector configuration W is the map

$$\begin{aligned} \chi \colon \{1, \ldots, n\}^d &\to \{-, 0, +\}\,, \\ (i_1, \ldots, i_d) &\mapsto \operatorname{sign}(\det(w^{i_1}, \ldots, w^{i_d})) \end{aligned}$$

which records for each d-tuple of vectors w^i if it forms a positively (or negatively) oriented basis of $\mathbb{R}^d$ or it is not a basis.

The *oriented matroid* of W is a combinatorial structure that can be given by any of the above data (co/vectors, co/circuits, or chirotopes) and defined/characterized in terms of any of the corresponding axiom systems. As an example, we state the chirotope/cocircuit translation, see Theorems 6.2.3 in [42] or 8.1.6 in [17].

Theorem 48. *Let $W \in \mathbb{R}^{d \times n}$ be a vector configuration with chirotope χ. Then the set of cocircuits is given by*

$$\mathcal{C}^*(\chi) = \Big\{ \big(\chi(I,1), \chi(I,2), \ldots, \chi(I,n)\big) \mid I \in \{1, \ldots, n\}^{d-1} \Big\}.$$

Conversely, let $W \in \mathbb{R}^{d \times n}$ be a vector configuration with cocircuits $\mathcal{C}^*$. Then there exists a unique pair of chirotopes $(\chi, -\chi)$ such that $\mathcal{C}^*(\chi) = \mathcal{C}^*(-\chi) = \mathcal{C}^*$.

The *face lattice* of $C = \operatorname{cone} W \subseteq \mathbb{R}^d$, the polyhedral cone generated by the vectors $w^1, \ldots, w^n$, can be obtained from the sign vectors of the linear subspace $\operatorname{im} W^\mathsf{T}$. In fact, it is the set $\operatorname{sign}(\operatorname{im} W^\mathsf{T})_\oplus = \operatorname{sign}(\operatorname{im} W^\mathsf{T}) \cap \{0, +\}^n$ with the partial order induced by the relation $+ > 0$. A face f of C corresponds to a supporting hyperplane with normal vector x such that $w^i \cdot x = 0$ for $w^i \in f$ and $w^i \cdot x > 0$ for $w^i \notin f$, lying on the positive side of the hyperplane. (The vector x lies on the corresponding face of the dual cone C^*.) Hence the face f with $I = \{i \mid w^i \in f\}$ is characterized by the sign vector $\tau = \operatorname{sign}(W^\mathsf{T} x) \in \operatorname{sign}(\operatorname{im} W^\mathsf{T})_\oplus$ with $I = \tau^0$. Moreover, for two faces f and f' of C with corresponding nonnegative sign vectors τ and τ', the order is reversed: $f \subseteq f'$ if and only if $\tau' \leq \tau$.

The *lineality space* of a cone C is given by the set $C \cap (-C)$. It is the minimal face of C, in the sense that it is contained in all faces. The lineality space of $C = \operatorname{cone} W$ is characterized by the maximal element of $\operatorname{sign}(\operatorname{im} W^\mathsf{T})_\oplus$ or, equivalently, by the maximal element of $\operatorname{sign}(\ker W)_\oplus$. Thereby, nonzero elements of $\operatorname{sign}(\ker W)_\oplus$ correspond to positive dependencies of vectors w^i (in the lineality space).

A cone C is called *pointed* if its lineality space is $\{0\}$, that is, if it has vertex 0. Note that, if $(+, \ldots, +)^\mathsf{T} \in \operatorname{sign}(\operatorname{im} W^\mathsf{T})_\oplus$ (that is, $\operatorname{sign}(\ker W)_\oplus = \{0\}$), then $C = \operatorname{cone} W$ is pointed.

Finally, we note that sign vectors of a linear subspace are closed under composition: Let S be a subspace of $\mathbb{R}^n$ and $\tau, \rho \in \operatorname{sign}(S)$. Then, also $\tau \circ \rho \in$

sign(S). To see this, let $u, v \in \mathbb{R}^n$ with $\tau = \text{sign}(u)$, $\rho = \text{sign}(v)$. Then, $\tau \circ \rho = \text{sign}(u + \varepsilon v) \in \text{sign}(S)$ for small $\varepsilon > 0$. Moreover, every nonzero sign vector of a linear subspace can be written as a conformal composition of sign vectors with minimal support, see Theorem 1 in [43], Proposition 5.35 in [4], or Theorem 3 in [40].

Theorem 49. *Let S be a subspace of $\mathbb{R}^n$ and $\tau \in \text{sign}(S)$ be nonzero. Then there are $\rho_i \in \text{sign}(S)$ with minimal support and $\rho_i \leq \tau$ such that*

$$\tau = \rho_1 \circ \cdots \circ \rho_N.$$

The ρ_i can be chosen such that $N \leq \min(\dim(S), |\text{supp}(\tau)|)$.

B A general theorem of the alternative

We recall a general theorem of the alternative for subspaces of $\mathbb{R}^n$ that allows to easily derive theorems of the alternative for sign vectors of a linear subspace and its orthogonal complement. For the relation to standard theorems of the alternative, see [36]; for the corresponding statements for arbitrary oriented matroids, see [7, Section 3.4] or [4, Chapter 5].

Definition 50. Let $x \in \mathbb{R}^n$, and let $I_1, \ldots, I_n$ be intervals of $\mathbb{R}$. We define the interval

$$\begin{aligned} I(x) &\equiv x_1 I_1 + \ldots + x_n I_n \\ &= \{x_1 y_1 + \ldots + x_n y_n \in \mathbb{R} \mid y_1 \in I_1, \ldots, y_n \in I_n\} \end{aligned}$$

and write $I(x) > 0$ if $y > 0$ for all $y \in I(x)$.

Theorem 51 (Theorem 22.6 in [44])**.** *Let S be a subspace of $\mathbb{R}^n$, and let $I_1, \ldots, I_n$ be intervals of $\mathbb{R}$. Then one and only one of the following alternatives holds:*

(a) There exists a vector $x = (x_1, \ldots, x_n)^\mathsf{T} \in S$ such that

$$x_1 \in I_1, \ \ldots, \ x_n \in I_n.$$

(b) There exists a vector $x^* = (x_1^*, \ldots, x_n^*)^\mathsf{T} \in S^\perp$ such that

$$x_1^* I_1 + \ldots + x_n^* I_n > 0.$$

Corollary 52. *Let S be a subspace of $\mathbb{R}^n$ and $\sigma \in \{-, 0, +\}^n$ be a nonzero sign vector. Then either (a) there exists a vector $x \in S$ with $x_i > 0$ for $i \in \sigma^+$ and $x_i < 0$ for $i \in \sigma^-$ or (b) there exists a nonzero vector $x^* \in S^\perp$ with $x_i^* \geq 0$ for $i \in \sigma^+$, $x_i^* \leq 0$ for $i \in \sigma^-$, and $x_i^* = 0$ otherwise. In terms of sign vectors, either there exists $\xi \in \text{sign}(S)$ with $\xi \geq \sigma$ or there exists a nonzero $\xi^* \in \text{sign}(S^\perp)$ with $\xi^* \leq \sigma$.*

Proof. By Theorem 51 with $I_i = (0, +\infty)$ for $i \in \sigma^+$, $I_i = (-\infty, 0)$ for $i \in \sigma^-$, and $I_i = (-\infty, +\infty)$ otherwise. □

Corollary 53. *Let S be a subspace of $\mathbb{R}^n$. Then,*

$$\operatorname{sign}(S^\perp) = \operatorname{sign}(S)^\perp.$$

Proof. ($\subseteq$): Let $\tau \in \operatorname{sign}(S^\perp)$ and $\rho \in \operatorname{sign}(S)$. Now, let $u \in S^\perp$ and $v \in S$ such that $\tau = \operatorname{sign}(u)$ and $\rho = \operatorname{sign}(v)$. Then, $u \cdot v = 0$ implies $\tau \cdot \rho = 0$, and hence $\tau \in \operatorname{sign}(S)^\perp$.

($\supseteq$): Let $\tau \notin \operatorname{sign}(S^\perp)$, that is, there exists no $x \in S^\perp$ such that $\operatorname{sign}(x) = \tau$. By Theorem 51 with $I_i = (0, +\infty)$ for $i \in \tau^+$, $I_i = (-\infty, 0)$ for $i \in \tau^-$, and $I_i = \{0\}$ otherwise, there exists a nonzero $x^* \in S$ such that $x^*_i \geq 0$ for $i \in \tau^+$ and $x^*_i \leq 0$ for $i \in \tau^-$. Let $\rho = \operatorname{sign}(x^*) \in \operatorname{sign}(S)$. Then, $\tau \cdot \rho \neq 0$, and hence $\tau \notin \operatorname{sign}(S)^\perp$. □

For an alternative proof, using Farkas Lemma, see Proposition 6.8 in [52].

References

[1] A. Albouy and Y. Fu. Some remarks about Descartes' rule of signs. *Elem. Math.*, 69, 2014.

[2] C. P. P. Arceo, E. C. Jose, A. R. Lao, and E. R. Mendoza. Reaction networks and kinetics of biochemical systems. *Math. Biosci.*, 283:13–29, 2017.

[3] C. P. P. Arceo, E. C. Jose, A. Marin-Sanguino, and E. R. Mendoza. Chemical reaction network approaches to biochemical systems theory. *Math. Biosci.*, 269:135–152, 2015.

[4] A. Bachem and W. Kern. *Linear programming duality.* Springer-Verlag, Berlin, 1992.

[5] S. Banach and S. Mazur. Über mehrdeutige stetige Abbildungen. *Stud. Math.*, 5:174–178, 1934.

[6] M. Banaji and C. Pantea. Some results on injectivity and multistationarity in chemical reaction networks. *SIAM J. Appl. Dyn. Syst.*, 15:807–869, 2016.

[7] A. Björner, M. Las Vergnas, B. Sturmfels, N. White, and G. M. Ziegler. *Oriented matroids*, volume 46 of *Encyclopedia Math. Appl.* Cambridge University Press, Cambridge, second edition, 1999.

[8] B. Boros and J. Hofbauer. Planar S-systems: Permanence. *J. Differential Equations*, 266:3787–3817, 2019.

[9] B. Boros, J. Hofbauer, and S. Müller. On global stability of the Lotka reactions with generalized mass-action kinetics. *Acta Appl. Math.*, 151:53–80, 2017.

[10] B. Boros, J. Hofbauer, S. Müller, and G. Regensburger. The center problem for the Lotka reactions with generalized mass-action kinetics. *Qual. Theory Dyn. Syst.*, 17:403–410, 2018.

[11] B. Boros, J. Hofbauer, S. Müller, and G. Regensburger. Planar S-systems: Global stability and the center problem. *Discrete Contin. Dyn. Syst. Ser. A*, 2(29):707–727, 2019.

[12] S. Chaiken. Oriented matroid pairs, theory and an electric application. In *Matroid theory (Seattle, WA, 1995)*, volume 197 of *Contemp. Math.*, pages 313–331. Amer. Math. Soc., Providence, RI, 1996.

[13] G. Craciun, A. Dickenstein, A. Shiu, and B. Sturmfels. Toric dynamical systems. *J. Symbolic Comput.*, 44:1551–1565, 2009.

[14] G. Craciun and M. Feinberg. Multiple equilibria in complex chemical reaction networks. I. The injectivity property. *SIAM J. Appl. Math.*, 65:1526–1546, 2005.

[15] G. Craciun, L. Garcia-Puente, and F. Sottile. Some geometrical aspects of control points for toric patches. In M. Dæhlen, M. S. Floater, T. Lyche, J.-L. Merrien, K. Morken, and L. L. Schumaker, editors, *Mathematical Methods for Curves and Surfaces*, volume 5862 of *Lecture Notes in Comput. Sci.*, pages 111–135, Heidelberg, 2010. Springer.

[16] G. Craciun, S. Müller, C. Pantea, and P. Y. Yu. A generalization of Birch's theorem and vertex-balanced steady states for generalized mass-action systems. 2018. arXiv:1802.06919 [math.DS].

[17] J. A. De Loera, J. Rambau, and F. Santos. *Triangulations. Structures for algorithms and applications.* Berlin: Springer, 2010.

[18] M. Feinberg. Complex balancing in general kinetic systems. *Arch. Rational Mech. Anal.*, 49:187–194, 1972/73.

[19] E. Feliu and C. Wiuf. Preclusion of switch behavior in networks with mass-action kinetics. *Appl. Math. Comput.*, 219:1449–1467, 2012.

[20] W. Fulton. *Introduction to toric varieties*, volume 131 of *Ann. of Math. Stud.* Princeton University Press, Princeton, NJ, 1993.

[21] I. M. Glazman and J. I. Ljubič. *Finite-dimensional linear analysis: a systematic presentation in problem form.* The M.I.T. Press, Cambridge, Mass.-London, 1974.

[22] G. Gnacadja. Univalent positive polynomial maps and the equilibrium state of chemical networks of reversible binding reactions. *Adv. in Appl. Math.*, 43:394–414, 2009.

[23] G. Gnacadja. A Jacobian criterion for the simultaneous injectivity on positive variables of linearly parameterized polynomial maps. *Linear Algebra Appl.*, 437:612–622, 2012.

[24] M. Gopalkrishnan, E. Miller, and A. Shiu. A geometric approach to the global attractor conjecture. *SIAM J. Appl. Dyn. Syst.*, 13:758–797, 2014.

[25] W. B. Gordon. On the diffeomorphisms of Euclidean space. *Am. Math. Mon.*, 79:755–759, 1972.

[26] J. Hadamard. Sur les transformations ponctuelles. *Bull. Soc. Math. Fr.*, 34:71–84, 1906.

[27] F. Horn. Necessary and sufficient conditions for complex balancing in chemical kinetics. *Arch. Rational Mech. Anal.*, 49:172–186, 1972/73.

[28] F. Horn and R. Jackson. General mass action kinetics. *Arch. Rational Mech. Anal.*, 47:81–116, 1972.

[29] G. J. O. Jameson. Counting zeros of generalised polynomials: Descartes' rule of signs and Laguerre's extensions. *Math. Gaz.*, 90:223–234, 2006.

[30] M. D. Johnston. Translated chemical reaction networks. *Bull. Math. Biol.*, 76:1081–1116, 2014.

[31] M. D. Johnston. A computational approach to steady state correspondence of regular and generalized mass action systems. *Bull. Math. Biol.*, 77:1065–1100, 2015.

[32] M. D. Johnston and E. Burton. Computing weakly reversible deficiency zero network translations using elementary flux modes. 2018. Submitted, arXiv:1808.09059 [math.OC].

[33] M. D. Johnston, S. Müller, and C. Pantea. A deficiency-based approach to parametrizing positive equilibria of biochemical reaction systems. *Bull. Math. Biol.*, 81:1143–1172, 2019.

[34] A. G. Khovanskiĭ. *Fewnomials*, volume 88 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, RI, 1991.

[35] E. N. Laguerre. Mémoire sur la théorie des équations numériques. *J. Math. Pures et Appl. (3)*, 9:99–146, 1883.

[36] G. J. Minty. A "from scratch" proof of a theorem of Rockafellar and Fulkerson. *Math. Program.*, 7:368–375, 1974.

[37] S. Müller, E. Feliu, G. Regensburger, C. Conradi, A. Shiu, and A. Dickenstein. Sign conditions for injectivity of generalized polynomial maps with applications to chemical reaction networks and real algebraic geometry. *Found. Comput. Math.*, 16:69–97, 2016.

[38] S. Müller and G. Regensburger. Generalized mass action systems: Complex balancing equilibria and sign vectors of the stoichiometric and kinetic-order subspaces. *SIAM J. Appl. Math.*, 72:1926–1947, 2012.

[39] S. Müller and G. Regensburger. Generalized mass-action systems and positive solutions of polynomial equations with real and symbolic exponents. In V. P. Gerdt, W. Koepf, E. W. Mayr, and E. H. Vorozhtsov, editors, *Computer Algebra in Scientific Computing. Proceedings of the 16th International Workshop (CASC 2014)*, volume 8660 of *Lecture Notes in Comput. Sci.*, pages 302–323, Cham, 2014. Springer.

[40] S. Müller and G. Regensburger. Elementary vectors and conformal sums in polyhedral geometry and their relevance for metabolic pathway analysis. *Front. Genet.*, 7(90):11 pages, 2016.

[41] L. Pachter and B. Sturmfels. Statistics. In *Algebraic statistics for computational biology*, pages 3–42. Cambridge Univ. Press, New York, 2005.

[42] J. Richter-Gebert and G. M. Ziegler. Oriented matroids. In *Handbook of discrete and computational geometry*, pages 111–132. CRC, Boca Raton, FL, 1997.

[43] R. T. Rockafellar. The elementary vectors of a subspace of R^N. In *Combinatorial Mathematics and its Applications (Proc. Conf., Univ. North Carolina, Chapel Hill, N.C., 1967)*, pages 104–127. Univ. North Carolina Press, Chapel Hill, N.C., 1969.

[44] R. T. Rockafellar. *Convex analysis.* Princeton University Press, Princeton, N.J., 1970.

[45] M. A. Savageau. Biochemical systems analysis. I. Some mathematical properties of the rate law for the component enzymatic reactions. *J. Theor. Biol.*, 25:365–369, 1969.

[46] F. Sottile. *Real Solutions to Equations from Geometry.* American Mathematical Society, Providence, RI, 2011.

[47] D. J. Struik, editor. *A source book in mathematics, 1200-1800.* Source Books in the History of the Sciences. Cambridge, Mass.: Harvard University Press, XIV, 1969.

[48] B. Sturmfels. *Solving systems of polynomial equations.* CBMS Regional Conf. Ser. in Math. Conference Board of the Mathematical Sciences, Washington, DC, 2002.

[49] D. A. S. J. Talabis, C. P. P. Arceo, and E. R. Mendoza. Positive equilibria of a class of power-law kinetics. *J. Math. Chem.*, 56:358–394, 2018.

[50] E. Tonello and M. D. Johnston. Network Translation and Steady-State Properties of Chemical Reaction Systems. *Bull. Math. Biol.*, 80:2306–2337, 2018.

[51] E. O. Voit. Biochemical systems theory: a review. *ISRN Biomath.*, 2013. Article ID 897658.

[52] G. M. Ziegler. *Lectures on polytopes.* Springer-Verlag, New York, 1995.