

# BIOCHEMICAL REACTION NETWORKS: AN INVITATION FOR ALGEBRAIC GEOMETERS

ALICIA DICKENSTEIN

ABSTRACT. This article is a survey of the recent use of some techniques from computational algebraic geometry to address mathematical challenges in systems biology. (Bio)chemical reaction networks define systems of ordinary differential equations with many parameters, which are needed for numerical simulations but that can be practically or provably impossible to identify. Under the standard modeling of mass-action kinetics, these equations depend polynomially on the concentrations of the chemical species. The algebraic theory of chemical reaction systems provides new tools to understand the dynamical behavior of (families of) chemical reaction systems by taking advantage of the inherent algebraic structure in the (parametric) kinetic equations.

## 1. INTRODUCTION

Chemical Reaction Network Theory (CNRT) has been developed over the last 40 years, initially through the work of Horn and Jackson and subsequently by Feinberg and his students and collaborators [24, 25, 26, 27, 28, 29, 30, 50, 51, 52, 53]. CRNT connected qualitative properties of ordinary differential equations corresponding to a reaction network to the network structure. In particular, assertions which are independent of specific parameter values have been obtained, in general assuming that all kinetics are of the mass-action form. New concepts were introduced, such as the deficiency of a reaction network, and several conditions were given on such networks for the existence, uniqueness, multiplicity and stability of fixed points. Foundational work has also been done by Vol'pert [79], with contributions with algebraic tools by Bykov, Kytmanov, Lazman and Yablonsky (see [8] and the references therein, together with more recent work as [44]).

The principal current application of these developments is in the realm of biochemical reaction networks, that is, chemical reaction networks in biochemistry. Systems biology's main goal is to understand the design principles of living systems. According to [36], the state of systems biology (at that moment, but still current) is like planetary astronomy science before Kepler and Newton and cannot be studied without mathematics and physics.

Recent work has focused on long-term dynamics as well as the capacity for multiple equilibria and how such dynamics depend on the specific rate parameters, mainly manipulating  $\mathbb{R}$ -linear combinations of the polynomials defining the dynamical systems (or equivalently, studying the kernel of the matrix  $M$  in (7)) [13, 15, 17, 76, 18, 42, 62, 72, 73, 74, 21].

---

AD was partially supported by UBACYT 20020100100242, CONICET PIP 20110100580, and ANPCyT PICT 2013-1110, Argentina.

We can use algebraic geometry to analyze systems biology models. Symbolic treatment of the parameters does not need a priori determination (which can be practically and theoretically impossible [17, 54]), as numerical simulations require. Karin Gatermann introduced the connection between mass-action kinetics and toric varieties at the beginning of the last decade [38, 39, 40]. Gunawardena also started approaching results from CRNT with algebraic tools over the last years [45, 46, 47, 61, 77, 78]. In joint work with Craciun, Shiu and Sturmfels, we studied in [12] *toric dynamical systems* (aka *complex balanced* mass-action systems) with an algebro-geometric perspective. The steady state locus of these systems coincides with the real points of a toric variety and, in appropriate coordinates in parameter space, the equations describing these complex balanced systems are also binomial. Advanced algebraic tools have been introduced by different authors over the last years [22, 16, 23, 32, 31, 34, 43, 48, 58, 60, 66, 69, 75].

Almost all cells in a body have the same genetic information. Multistationarity (see Definition 2.4) provides a mechanism for switching between different response states in cell signaling systems and enables multiple outcomes for cellular-decision making [59]. Questions about steady states in biochemical reaction networks under mass-action kinetics are fundamentally questions about (nonnegative) real solutions to parametrized polynomial ideals.

We present in Section 2 the basic notations and concepts about chemical reaction networks. Section 3 concentrates on the important enzymatic networks, that we use to exemplify questions on multistationarity. Section 4 is devoted to the notion of steady state invariants. Invariants depending on selected variables can be used to understand the design of the different mechanisms. We distinguish four levels of invariants and we show applications to model selection, to study absolute concentration robustness and to obtain nontrivial bounds via implicit dose-response curves. It follows that the study of ideals over polynomial rings unveils features of the steady states not visible working only with coefficients in  $\mathbb{R}$ , but further tools from real algebraic geometry are required. Finally, in Section 5 we summarize recent general results on sign conditions for multistationarity, that hold beyond the framework of chemical reaction networks. Along the text, we recall results from joint recent papers and preprints [58, 65, 67, 68, 69]. A more comprehensive account will appear in the book in progress with Elisenda Feliu [19].

We end this introduction with pointers to a few important subjects we have not addressed in this text, together with an overview of general goals for our approach and new algebro-geometric tools that we expect to incorporate.

All biological processes are complex and involve many variables and (unknown) reaction rate constants. An apparent solution to the complexity challenges in cellular networks consists of studying smaller subunits that one can analyze separately. In fact, essential qualitative features of biological processes can usually be understood or qualitatively approximated for parameters in a certain range, in terms of a small number of crucial variables [59]. In [71], the authors defined *network motifs* as patterns of interconnections that recur in many different parts of a big network. Study of subnetworks to determine multistationarity has been addressed for instance in [9] (via *elementary flux modes*) and [35, 57] (via versions of the implicit function theorem). We expect that tools from deformation theory could help extending these results to the case of degenerate steady states.

Differential algebra methods and in particular differential elimination methods, provide tools for searching hidden relations which are consequences of our differential-algebraic polynomial (nonlinear) equations. They have been used for parameter estimation in nonlinear dynamical systems and model reduction of biochemical systems (via implicit *quasi-steady state approximation*) and some related software is available [3, 4, 5, 6, 7]. It would be interesting to further explore the use of these tools.

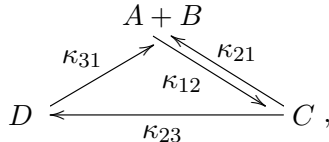
We have not discussed the global dynamic behaviour of the systems. The main open conjecture in the field of Chemical Reaction Network Theory is the *Global Attractor Conjecture*, which dates back to the early 1970s. Complex balanced chemical reaction networks associated with weakly reversible graphs, possess a unique positive steady state in any given stoichiometric compatibility class (see Section 2), which shows local asymptotic stability deduced from the existence of a Lyapunov function. The Global Attractor Conjecture asserts that this is in fact a global attractor for the dynamics. This statement is proven in the absence of steady states with zero coordinates, in case the reaction graph is connected or in case the dynamics occurs in dimension at most three, but the combinatorics of zero coordinates of the boundary steady states makes the search for a proof of the general result highly complicated [1, 12, 16, 55]. At the time of the revision of this article, G. Craciun has posted a first version of an article which would positively solve the Conjecture [11].

Tools from elimination theory in computational algebraic geometry and from real algebraic geometry can be used to study the number and stability of steady states in families, as well as the possible occurrence of bifurcations and oscillations in polynomial (nonlinear) dynamical systems. One general goal is to partition the positive orthant in constant rate space of a given biochemical network into semialgebraic sets, in such a way that on each chamber the dynamic behaviour can be determined. The study of properties that depend on the structure of the network and are independent of the particular reaction rate constants in this semialgebraic decomposition of parameter space, would allow us to see “the woods” and not “only the trees”. The super goal is to understand the basic mechanisms in nature for multistationarity and for oscillations. In theory, computational algebraic geometry can give many answers. In practice, these responses tend to be too complex to be understood or computed. Many answers are missing and require the combination of tools from computer algebra, real algebraic geometry, numerical algebraic geometry, discrete mathematics, dynamical systems, and biochemistry!

## 2. BASICS ON CHEMICAL REACTION NETWORKS (CRN)

We start with a simple but meaningful example of a biochemical reaction network: the T-cell signal transduction model proposed by the immunologist McKeithan [63]. The main task of the immune system is to recognize that a strange body has entered the organism. T-cell receptors bind to both self-antigens and foreign antigens and the dynamical features of this model give a possible explanation of how T-cells can be sensitive and specific in recognizing self versus foreign antigens. A mathematical study of the dynamics of this network was done

by Sontag in [76]. In its simplest case, the network of reactions is as follows:



where  $A$  denotes the T-cell receptor protein,  $B$  denotes the Major Histocompatibility protein Complex (MHC) of antigen-presenting cell,  $C$  denotes the biochemical species  $A$  bound to species  $B$ , and  $D$  denotes an activated (phosphorylated) form of  $C$ . The binding of  $A$  and  $B$  forms  $C$ , which undergoes a modification into its activated form  $D$  before “transmitting a signal” (that is, before participating in another chemical reaction). The general mechanism proposed by McKeithan includes several activated forms of  $C$ , until a final active form that “triggers the attack” to the foreign antigen is obtained.

This biochemical reaction network has:

- $r = 4$  reactions among
- $m = 3$  complexes  $A + B$ ,  $C$ , and  $D$ , which are composed by
- $s = 4$  species  $A$ ,  $B$ ,  $C$ ,  $D$ , and
- $r = 4$  reaction rate constants  $\kappa_{12}, \kappa_{21}, \kappa_{23}, \kappa_{31} \in \mathbb{R}_{>0}$  for the different reactions.

A kinetics is then attached to this labeled directed graph to describe how the concentrations  $x_A, x_B, x_C, x_D$  of the different biochemical species evolve in time.

McKeithan assumes that the vector of concentrations  $x(t) = (x_A(t), x_B(t), x_C(t), x_D(t))$  evolves according to mass-action kinetics, which is a modeling commonly used in chemistry and biology when there are sufficiently many molecules that are well mixed. The *Law of Mass Action* was proposed by two Norwegians: Peter Waage (1833–1900), a chemist, and Cato Guldberg (1836–1902), a mathematician, in an article published in Norwegian in 1864. Their work was then published in French in 1867 and finally, a fuller and further developed account appeared in German in 1879, and was then recognized (in the meantime this principle was rediscovered by van’t Hoff). The Law of Mass Action is derived from the idea that the the rate of an elementary reaction is proportional to the probability of collision of reactants (under an independence assumption), that is, to the product of their concentrations. We write the precise formulation in (1) below.

The explicit differential equations for the concentrations  $x(t)$  in the T-cell signal transduction model are the following:

$$\begin{array}{llll}
 \frac{dx_A}{dt} & = & -\kappa_{12}x_Ax_B + \kappa_{21}x_C + \kappa_{31}x_D & = & -\kappa_{12}x^{(1,1,0,0)} + \kappa_{21}x^{(0,0,1,0)} + \kappa_{31}x^{(0,0,0,1)} \\
 \frac{dx_B}{dt} & = & -\kappa_{12}x_Ax_B + \kappa_{21}x_C + \kappa_{31}x_D & = & -\kappa_{12}x^{(1,1,0,0)} + \kappa_{21}x^{(0,0,1,0)} + \kappa_{31}x^{(0,0,0,1)} \\
 \frac{dx_C}{dt} & = & \kappa_{12}x_Ax_B - \kappa_{21}x_C - \kappa_{23}x_C & = & \kappa_{12}x^{(1,1,0,0)} - (\kappa_{21} + \kappa_{23})x^{(0,0,1,0)} \\
 \frac{dx_D}{dt} & = & \kappa_{23}x_C - \kappa_{31}x_D & = & \kappa_{23}x^{(0,0,1,0)} - \kappa_{31}x^{(0,0,0,1)}.
 \end{array}$$

In general, the starting data for a chemical reaction network are a finite set of  $s$  species (whose concentrations  $x_1, \dots, x_s$  will be our variables), a finite set of  $r$  reactions (labeled edges  $i \xrightarrow{\kappa_{ij}} j$ , where  $\kappa_{ij} \in \mathbb{R}_{>0}$  are the reaction rate constants), between  $m$  complexes  $y_1, \dots, y_m \in \mathbb{Z}_{\geq 0}^s$  among the species (which are classically represented as nonnegative integer combinations of the species and which give rise to monomials in the concentrations

of the chemical species  $x^{y_i} = x_1^{y_{i1}} x_2^{y_{i2}} \cdots x_s^{y_{is}}$ . The entries of the complexes are called stoichiometric coefficients.

**Definition 2.1.** A *chemical reaction network* (CRN) is a finite directed graph

$$G = (V, E, (\kappa_{ij})_{(i,j) \in E}, (y_i)_{i=1, \dots, m})$$

whose vertices are labeled by complexes and whose edges are labeled by positive real numbers. Mass-action kinetics specified by the network  $G$  gives the following autonomous system of ordinary differential equations in the concentrations  $x_1, x_2, \dots, x_s$  of the species as *functions of time  $t$* :

$$(1) \quad \frac{dx}{dt} = \sum_{(i,j) \in E} \kappa_{i,j} x^{y_i} (y_j - y_i).$$

Note that system (1) is of the form

$$(2) \quad \frac{dx_k}{dt} = f_k(x), \quad k = 1, \dots, s,$$

where  $f_1, \dots, f_s$  are polynomials in  $\mathbb{R}[x_1, \dots, x_s]$ .

A first natural question is which autonomous polynomial dynamical systems come from a CRN under mass-action kinetics. The answer is due to Hárs and Tóth:

**Lemma 2.2** ([49]). *A polynomial dynamical system  $dx/dt = f(x)$  in  $s$  variables  $(x_1, \dots, x_s)$  arises from a CRN under mass-action kinetics if and only if there exists real polynomials  $p_k, q_k, k = 1, \dots, s$ , with non negative coefficients such that  $f_k = p_k - x_k q_k$  for all  $k$ .*

The necessary condition that each monomial with negative coefficient in the polynomial  $f_k$  has to be divisible by  $x_k$  is straightforward from (1). The converse is constructive. One interesting feature that follows from this constructive proof is the fact that the polynomials  $f_k$  do not determine the network, only (almost) the source complexes of the reactions (those labeling the initial node of a directed edge). We refer the reader to [17, 54] for extensions and precisions of Lemma 2.2, in particular, identifiability of the reaction rate constants  $\kappa_{ij}$  for a given network. In general, one assumes the structure of the reaction network and would like to infer dynamical properties of the system from this structure, even if most reaction rate constants are unknown.

The restriction on the coefficients of a CRN under mass-action kinetics given by Lemma 2.2 is satisfied for instance by the oscillatory Lotka-Volterra equations, but not by the “chaotic” Lorenz equations

$$\frac{dx_1}{dt} = \alpha x_2 - \alpha x_1, \quad \frac{dx_2}{dt} = \gamma x_1 - x_2 - x_1 x_3, \quad \frac{dz}{dt} = x_1 x_2 - \beta x_3, \quad \alpha, \beta, \gamma \in \mathbb{R}_{>0},$$

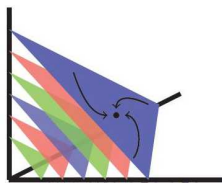
due to the existence of the term  $-x_1 x_3$  in  $f_2$ .

**Definition 2.3.** The steady state variety  $V(f)$  of the kinetic system (2) equals the nonnegative real zeros of  $f_1, \dots, f_s$ , that is, the nonnegative points of the real algebraic variety cut out by  $f_1, \dots, f_s$ . Any element of  $V(f)$  is called a steady state of the system.

Note that the positive solutions of the system  $x_1 f_1 = \cdots = x_s f_s = 0$  equal the positive solutions of  $f_1 = \cdots = f_s = 0$  (but of course the dynamics of the corresponding differential systems is different). So, any system of  $s$  real polynomials in  $s$  variables defines the positive

steady states of a CRN under mass-action kinetics. However, realistic models have particular features that produce interesting results. We will focus in particular on enzymatic networks.

Another direct consequence of the form of the equations in (1) is that for any trajectory  $x(t)$ , the vector  $\frac{dx}{dt}$  lies for all  $t$  in the so called *stoichiometric subspace*  $S$ , which is the linear subspace generated by the differences  $\{y_j - y_i \mid (i, j) \in E\}$ . Using the shape of the polynomials  $f_k = p_k - x_k q_k$  in Lemma 2.2, it is straightforward to see that a trajectory  $x(t)$  starting at a nonnegative point  $x(0)$  lies in the closed polyhedron  $(x(0) + S) \cap \mathbb{R}_{\geq 0}^s$  for all  $t \geq 0$ , called a *stoichiometric compatibility class*. The (linear) equations of  $x(0) + S$  are called conservation relations.



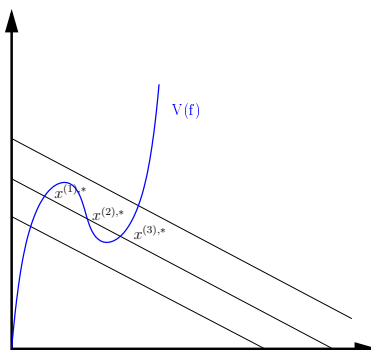
Different stoichiometric compatibility classes

Note that for any autonomous dynamical system of the form (2), any linear relation  $\sum_{i=1}^s c_i f_i = 0$  with real coefficients  $c_1, \dots, c_s$ , gives rise to the restriction that  $\sum_{i=1}^s c_i x_i$  has to be constant along trajectories. In our setting, the linear equations for  $S$  give conservation relations, but for specific  $f_1, \dots, f_s$  there could be further linear constraints.

As we pointed out in the introduction, a central notion is the following:

**Definition 2.4.** We say that system (1) exhibits multistationarity if there exist at least two steady states in the same stoichiometric compatibility class.

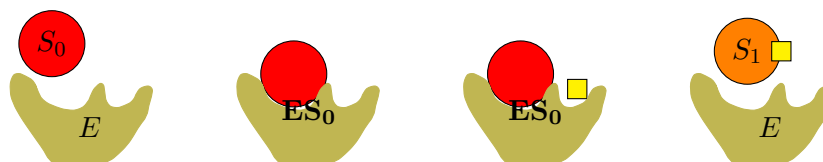
The following figure illustrates the intersection of the steady state variety  $V(f)$  with different stoichiometric compatibility classes. The middle one has 3 different steady states  $x^{(1),*}, x^{(2),*}, x^{(3),*}$ , so the system exhibits multistationarity.



In the following section we will concentrate on multistationarity questions of enzymatic networks. Section 5 presents recent general mathematical results to preclude or allow the occurrence of multistationarity based on sign vectors.

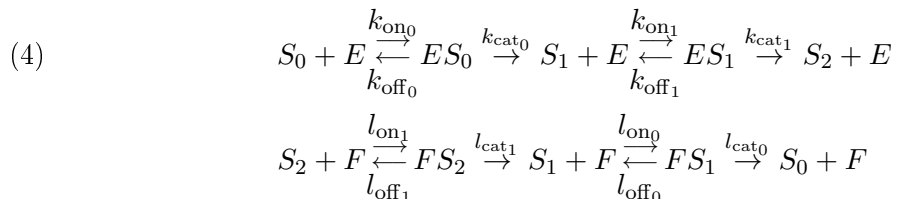
### 3. ENZYMATIC NETWORKS

The Nobel Prize in Physiology or Medicine 1992 was awarded jointly to Edmond H. Fischer and Edwin G. Krebs “for their discoveries concerning reversible protein phosphorylation as a biological regulatory mechanism”. Phosphorylation/dephosphorylation are post-translational modification of proteins mediated by enzymes, particular proteins that add or take off a phosphate group at a specific site, inducing a conformational change that allows/prevents the protein to perform its function. The standard building block in cell signaling is the following enzyme mechanism, which is called a Michaelis-Menten mechanism, named after the German biochemist Leonor Michaelis and the Canadian physician Maud Menten. This basic network involves four species: the substrate  $S_0$ , the phosphorylated substrate  $S_1$ , the enzyme  $E$  and the intermediate species  $ES_0$ . The enzyme  $E$  is not “consumed” after the whole mechanism, which is assumed to be with mass-action kinetics. The concentration of the donor of the phosphate group is considered to be constant, thus hidden in the reaction rate constants and ignored. A scheme is as follows, with the 3 reaction rate constants called  $k_{on}$ ,  $k_{off}$ ,  $k_{cat}$ :



One canonical class of biological systems exhibiting multistationarity are protein kinase mechanisms that involve multiple phosphorylation of a substrate. There are (substrate) proteins in humans that are known to have more than 150 possible phosphorylation sites [78].

The following CRN corresponds to the case of  $n = 2$  *sequential* phosphorylations:



This network involves nine species: the substrates with zero, one and two phosphorylated sites  $S_0, S_1, S_2$  (known as phosphoforms), the intermediate species  $ES_0, ES_1, FS_1, FS_2$  plus two enzymes  $E, F$  ( $E$  is called a kinase and  $F$  a phosphatase), and ten complexes denoted as integer linear combinations of species by  $S_0 + E, S_1 + E, S_2 + E, ES_0, ES_1, S_0 + F, S_1 + F, S_2 + F, FS_1, FS_2$ . Renaming the variables and the complexes following the previous ordering, we

get the following dynamical system for the concentrations under mass-action kinetics:

$$\begin{aligned}
\frac{dx_1}{dt} &= -k_{\text{on}_0}x_1x_8 + k_{\text{off}_0}x_4 + l_{\text{cat}_0}x_6 & \frac{dx_6}{dt} &= l_{\text{on}_0}x_2x_9 - (l_{\text{cat}_0} + l_{\text{off}_0})x_6 \\
\frac{dx_2}{dt} &= -k_{\text{on}_1}x_2x_8 + k_{\text{cat}_0}x_4 + k_{\text{off}_1}x_5 & \frac{dx_7}{dt} &= l_{\text{on}_1}x_3x_9 - (l_{\text{cat}_1} + l_{\text{off}_1})x_7 \\
& -l_{\text{on}_0}x_2x_9 + l_{\text{off}_0}x_6 + l_{\text{cat}_1}x_7 & \frac{dx_8}{dt} &= -k_{\text{on}_0}x_1x_8 - k_{\text{on}_1}x_2x_8 + (k_{\text{off}_0} + k_{\text{cat}_0})x_4 \\
& & & + (k_{\text{off}_1} + k_{\text{cat}_1})x_5 \\
\frac{dx_3}{dt} &= k_{\text{cat}_1}x_5 - l_{\text{on}_1}x_3x_9 + l_{\text{off}_1}x_7 & & \\
\frac{dx_4}{dt} &= k_{\text{on}_0}x_1x_8 - (k_{\text{off}_0} + k_{\text{cat}_0})x_4 & \frac{dx_9}{dt} &= -l_{\text{on}_0}x_2x_9 - l_{\text{on}_1}x_3x_9 + (l_{\text{cat}_0} + l_{\text{off}_0})x_6 \\
\frac{dx_5}{dt} &= k_{\text{on}_1}x_2x_8 - (k_{\text{off}_1} + k_{\text{cat}_1})x_5 & & + (l_{\text{cat}_1} + l_{\text{off}_1})x_7
\end{aligned}$$

The stoichiometric subspace  $S$  has codimension 3, so there are 3 linearly independent conservation relations, usually taken as total substrate, total kinase and total phosphatase:

$$\begin{aligned}
(5) \quad & x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 = S_{\text{tot}} \\
& x_4 + x_5 + x_8 = E_{\text{tot}} \\
& x_6 + x_7 + x_9 = F_{\text{tot}}.
\end{aligned}$$

So, there are only 6 linearly independent differential equations in the system. The constants  $(S_{\text{tot}}, E_{\text{tot}}, F_{\text{tot}})$  are determined by the initial conditions. We see that each stoichiometric compatibility class is compact since adding the 3 conservation relations we get a positive linear combination involving all the variables equal to a positive number, so each class is bounded (and closed). In general, the  $n$ -site phosphorylation system is of great biochemical importance: it is a recurring network motif in many networks describing biochemical processes.

The common zeros of  $f_1, \dots, f_s$  equal the common zeros of the ideal of their polynomial consequences (the *steady state ideal*):

$$(6) \quad I_f = \{g_1f_1 + \dots + g_sf_s : g_i \in \mathbb{R}[x_1, \dots, x_s], i = 1, \dots, s\}.$$

The polynomials  $f_1, \dots, f_s$  are generators of  $I_f$ . We refer the reader to [10] for the basic notions of polynomial ideals and Gröbner bases.

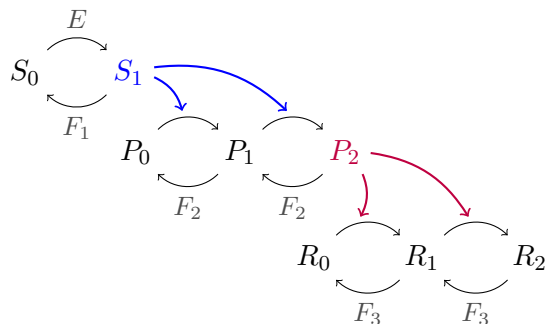
If the steady state ideal  $I_f$  is a binomial ideal, that is, if it can be generated by polynomials which are binomials (i.e., polynomials with two terms), we say that the system has *toric steady states*. We prove in [69] that the chemical reaction system associated with the multisite  $n$ -phosphorylation of a protein by a kinase/phosphatase pair in a sequential and distributive mechanism (with the same structure as the mechanism in (4) but with  $n$  sites) has *toric steady states*. This result was implicit in [80] and it is a particular case of [77]. This system has  $3n + 3$  species,  $4n + 2$  complexes and still 3 linearly independent conservation relations.

Wang and Sontag studied in [80] the number of steady states in the general  $n$ -site sequential distributive phosphorylation network and showed that there are at most  $2n - 1$  steady states in each stoichiometric compatibility class. They also identified a particular open set



in the positive orthant of the constant rate space  $\mathbb{R}_{>0}^r$  where the number of positive steady states in the same compatibility class is  $n + 1$  (for  $n$  even) or  $n$  (for  $n$  odd) and conjectured that the maximum possible number is  $n + 1$  for any  $n$ . It was shown in [37], that in fact for any  $n = 3$  and 4 there can be up to  $2n - 1$  stoichiometrically compatible positive steady states, for particular choices of the reaction rate constants. Even for  $n = 3$  it is very complicated to give a precise description of the (semialgebraic) regions in which  $\mathbb{R}_{>0}^r$  can be partitioned according to the maximal number of steady states and no study explains for the moment how many of the known steady states in a given compatibility class are *attractors* for the dynamics.

We show in [68] that many other important networks have toric steady states including most of the motifs of enzyme cascades studied in [33], for example, the following cascade of phosphorylations known as the MAPK/ERK pathway:



Each curved arrow in this diagram represents a digraph with 3 nodes as in (3), where the enzyme is the label of the arrow. Note that the phosphorylated (or double phosphorylated) substrate in the upper reactions acts as an enzyme down the cascade.

In general, deciding multistationarity amounts to the difficult question of determining emptiness of a (complicated) semialgebraic set, which is in principle algorithmic but unfeasible in practice. For chemical reaction systems with toric steady states for every choice of positive reaction rate constants, we have the following explicit criterion for multistationarity [69]. First, if the system has toric steady states for every choice of positive reaction rate constants, the steady states can be *explicitly parametrized* by monomials (or shown to be empty). That is, we can check for non emptiness and the positive steady states can be parametrized by a *monomial map*

$$t \mapsto (c_1(\kappa)t^{v_1}, \dots, c_s(\kappa)t^{v_s}),$$

where  $t \in \mathbb{R}_+^d$ ,  $d$  is the dimension of the steady state variety and  $c_1, \dots, c_s$  are rational functions of the  $\kappa_{ij}$ . Now, we can check for multistationarity in an algorithmic way (under the conditions detailed in [69, Section 3]). Call  $V \in \mathbb{N}^{d \times s}$  the matrix with columns  $v_1, \dots, v_s$ .

The following is a simplified version of Theorem 5.5 in [69]:

**Theorem 3.1** ([69]). *Fix a chemical reaction network  $G$  with  $s$  species, under mass-action kinetics such that there exist positive constants  $\mu_{ij}$  for all reactions such that  $\sum \mu_{ij}(y_j - y_i) = 0$ .<sup>1</sup> Assume the system has toric steady states for all reaction rate constants and it satisfies Condition 3.1 or 3.16 in [69]. Let  $V \in \mathbb{N}^{d \times s}$  be a matrix giving the exponents of*

<sup>1</sup>Note that by (1), this condition is necessary for the existence of a positive steady state

a parametrization of the positive steady state variety. There exists a reaction rate constant vector such that the resulting chemical reaction network exhibits two different positive steady states in the same stoichiometric compatibility class if and only if there exists an orthant  $O$  of  $\mathbb{R}^s$  of any positive dimension that the two intersections  $O \cap \text{image}(V)$  and  $O \cap S$  are both non empty, or in other words, if and only if there exist non-zero  $\alpha \in \text{image}(V)$  and  $\beta$  in the stoichiometric subspace  $S$  with  $\text{sign}(\alpha_i) = \text{sign}(\beta_i)$  for all  $i = 1, \dots, s$ .

We will present a recent general result on multistationarity in Section 5.

#### 4. STEADY STATE INVARIANTS

We keep in this section the notations of Section 2. Note that we can also write the polynomial autonomous system (1) which models the kinetics of a chemical reaction network, as a real matrix  $M \in \mathbb{R}^{s \times m}$  multiplied by a vector of monomials  $\Psi(x)$  with  $i$ -th coordinate equal to  $x^{y_i}$ :

$$(7) \quad \frac{dx}{dt} = f(x) = M(\Psi(x)).$$

**Definition 4.1.** A steady state invariant (or simply, an invariant) is a polynomial that vanishes on the steady state variety  $V(f)$ .

The given polynomials  $f_1, \dots, f_s$  are trivially steady state invariants. But we are interested in describing new invariants that reveal further properties of the system. In many cases, it is most important to find invariants that only depend on a selected subset of variables, which usually correspond to those concentrations that are easier to measure, or to concentrations one wants to relate at steady state.

We can distinguish four “levels” of invariants.

- Level 1: Any element of the rowspan of  $M$  defines an invariant which is an  $\mathbb{R}$ -linear combination of  $f_1, \dots, f_s$ . Level 1 invariants depending on fewer complexes can be simply obtained by Gaussian elimination. For any element  $\lambda$  in the rowspan of  $M$ , the sum  $\sum_{i=1}^m \lambda_i v_i$  vanishes for any vector  $v \in \ker(M)$ ; in particular, the polynomial  $\sum_{i=1}^m \lambda_i x^{y_i}$  vanishes at steady state.
- Level 2: Any polynomial in the steady state ideal  $I_f \subset \mathbb{R}[x_1, \dots, x_s]$  defined in (6) is an invariant, which can be obtained via computational algebraic geometry methods as a polynomial linear combination of  $f_1, \dots, f_s$ . In particular, any invariant of Level 1 is an invariant of Level 2, and the inclusion is strict. Note that any invariant of Level 2 vanishes on all *complex* common zeros of  $f_1, \dots, f_s$ . Elimination ideals  $I_f \cap \mathbb{R}[x_i, i \in \Gamma]$  for a given subset  $\Gamma$  of  $\{1, \dots, s\}$ , can be effectively computed with Gröbner basis methods, which are for instance efficiently implemented in the free computer algebra systems Singular [20] or Macaulay2 [41]. We will mainly deal with positive steady states, which in particular have nonzero coordinates. Primary decomposition of ideals has been applied in [75] to describe boundary steady states (with some zero coordinate).
- Level 3: Any polynomial in the *radical*  $\sqrt{I_f}$  of the ideal  $I_f$  is an invariant. By Hilbert Nullstellensatz, these are precisely those polynomials that vanish on all the complex common zeros of  $f_1, \dots, f_s$ . The radical ideal  $\sqrt{I_f}$  can also be computed via computational algebraic geometry, keeping the same zeros but without “multiplicity”.

Level 4: Any polynomial which vanishes on  $V(f)$ , that is on the nonnegative real zeros of  $I_f$ , is an invariant by definition. These polynomials form an ideal  $\mathbb{R}_{\geq 0}\sqrt{I_f}$  that we could call the *positive real radical* of  $I_f$ . The positive real radical is in turn contained in the real radical  $\sqrt{I_f}$  of  $I_f$  composed of all polynomials which vanish on the real zeros of  $I_f$ . These notions pertain to the (difficult) realm of real algebraic geometry.

In general, we have that

$$(8) \quad I_f \subset \sqrt{I_f} \subset \sqrt[3]{I_f} \subset \mathbb{R}_{\geq 0}\sqrt{I_f},$$

and the inclusions are in general strict. A simple example for  $n = 1$  is given by the ideal  $I_f \subset \mathbb{R}[x]$  generated by the polynomial  $f = x^2(x^2 - 1)(x^2 + 1)$ . In this case,  $x(x^2 - 1)(x^2 + 1)$  lies in  $\sqrt{I_f} \setminus I_f$ ,  $x(x^2 - 1)$  lies in  $\sqrt[3]{I_f} \setminus \sqrt{I_f}$ , the polynomial  $x(x - 1)$  lies in  $\mathbb{R}_{\geq 0}\sqrt{I_f} \setminus \sqrt[3]{I_f}$  and  $x - 1$  vanishes on the positive real zeros of  $I_f$ . Another simple example in one variable shows that algebraic extensions enter into the picture: for instance, take  $f = x^5 - 2x$ ; then  $\mathbb{R}_{\geq 0}\sqrt{I_f}$  can be generated by the polynomial  $x^2 - \alpha$ , with the additional information that  $\alpha^2 - 2 = 0$  and  $\alpha > 0$ . However, the containments in (8) are equalities for the most usual enzymatic networks.

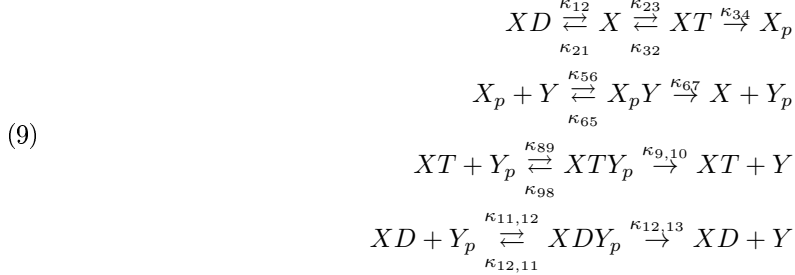
Invariants can be used to check the (un)correctness of a proposed model [61]. A baby example of this application taken from [46] is the following. In the sequential enzymatic mechanism for  $n = 2$ , we get by elimination of variables an invariant of Level 2 of the form  $x_j(x_1x_3 - Kx_2^2)$  with  $j = 8$  or  $j = 9$ , where  $K$  depends on the (unknown) reaction rate constants but not on the initial conditions! Recall that  $x_1, x_2, x_3$  denote the concentrations at steady state of the unphosphorylated, singly phosphorylated or doubly phosphorylated substrate and  $x_8, x_9$  denote the enzymes, which can be measured and are assumed to be positive. So the “values at steady state”  $(x_1, x_2, x_3)$  of the concentrations for different runs should satisfy, according to this model, that the points  $(x_1x_3, x_2^2)$  lie (approximately) on a line. Even if the slope is unknown, plotting these points allows to check the correctness of the model. In fact,  $K$  is the following explicit rational function in the reaction rate constants (obtained via elimination in the polynomial ring with variables  $x_i$  and  $\kappa_{ij}$ ): it is the quotient  $P_1/Q_1$  of the following polynomials:

$$\begin{aligned} P_1 &= \kappa_{10,7}\kappa_{25}\kappa_{41}\kappa_{54}\kappa_{79}\kappa_{96} + \kappa_{10,7}\kappa_{25}\kappa_{42}\kappa_{54}\kappa_{79}\kappa_{96} + \kappa_{10,8}\kappa_{25}\kappa_{41}\kappa_{54}\kappa_{79}\kappa_{96} + \kappa_{10,8}\kappa_{25}\kappa_{42}\kappa_{54}\kappa_{79}\kappa_{96}, \\ Q_1 &= \kappa_{10,7}\kappa_{14}\kappa_{42}\kappa_{52}\kappa_{8,10}\kappa_{96} + \kappa_{10,7}\kappa_{14}\kappa_{42}\kappa_{52}\kappa_{8,10}\kappa_{97} + \kappa_{10,7}\kappa_{14}\kappa_{42}\kappa_{53}\kappa_{8,10}\kappa_{96} \\ &\quad + \kappa_{10,7}\kappa_{14}\kappa_{42}\kappa_{53}\kappa_{8,10}\kappa_{97}. \end{aligned}$$

Note that steady states correspond to nonnegative constant solutions of (1), so if for any finite  $t$  the system is at steady state, then the trajectory is constant. If there exists a positive conservation relation  $\sum_{i=1}^s c_i \frac{dx_i}{dt} = 0$  with  $c_1, \dots, c_s > 0$  as in this example, the trajectories are bounded (so the system is *conservative*) and then each trajectory is defined for any  $t \geq 0$ . The “values at steady state” are the limit values  $\lim_{t \rightarrow \infty} x_i(t)$  (when these limits exist), which can be approximated with experimental measurements.

**4.1. Invariants and the notion of Absolute Concentration Robustness.** Shinar and Feinberg introduced in [74] the notion of Absolute Concentration Robustness (ACR, for short) of a given chemical species  $x_j$ . This happens when the  $j$ -th coordinate of the positive steady states of the system have a fixed value, independent of the given positive steady state and even independent of the value of the conservation relations (see also [72] for the notion of robustness of the output with respect the initial conditions). This is a very peculiar

feature that shows up in real examples. Here is a particular mechanism extracted from [74]. The enzyme  $X$  is a kinase (known as EnvZ) present in the the bacteria Escherichia Coli, that can be self-transformed into  $XD$  and  $XT$  and it can then be self-phosphorylated to produce the species  $X_p$ . In  $X_p$  form, it can react with species  $Y$  (known as OmpR) to obtain the phosphorylated form  $Y_p$ , while  $XD$  and  $XT$  can dephosphorylate  $Y_p$  by the standard Michaelis-Menten mechanism:



We denote by  $x_1, \dots, x_9$  the species concentrations as follows:  $x_{XD} = x_1$ ,  $x_X = x_2$ ,  $x_{XT} = x_3$ ,  $x_{X_p} = x_4$ ,  $x_Y = x_5$ ,  $x_{X_p Y} = x_6$ ,  $x_{Y_p} = x_7$ ,  $x_{XTY_p} = x_8$ ,  $x_{XDY_p} = x_9$ .

In fact, this system has toric steady states. Indeed, an ideal is binomial if and only if any *reduced Gröbner basis* is composed of binomials. Any such basis gives a binomial system of generators for the ideal. The reduced Gröbner basis of  $I_f$  with respect to the lexicographical order  $x_1 > x_2 > x_4 > x_5 > x_6 > x_8 > x_9 > x_3 > x_7$  consists of the following binomials:

$$\begin{aligned}
 g_1 &= [\kappa_{89}\kappa_{12}\kappa_{23}\kappa_{9,10}(\kappa_{12,11} + \kappa_{12,13}) + \kappa_{11,12}\kappa_{21}\kappa_{12,13}(\kappa_{98} + \kappa_{9,10})(\kappa_{32} + \kappa_{34})]x_3x_7 - \\
 &\quad - [\kappa_{23}\kappa_{34}\kappa_{12}(\kappa_{12,11} + \kappa_{12,13})(\kappa_{98} + \kappa_{9,10})]x_3 \\
 g_2 &= [-\kappa_{11,12}\kappa_{21}\kappa_{34}(\kappa_{98} + \kappa_{9,10})(\kappa_{32} + \kappa_{34})]x_3 + \\
 &\quad + [\kappa_{11,12}\kappa_{21}\kappa_{12,13}(\kappa_{98} + \kappa_{9,10})(\kappa_{32} + \kappa_{34}) + \kappa_{12}\kappa_{23}\kappa_{89}\kappa_{9,10}(\kappa_{12,11} + \kappa_{12,13})]x_9 \\
 g_3 &= [-\kappa_{23}\kappa_{34}\kappa_{89}\kappa_{12}(\kappa_{12,11} + \kappa_{12,13})]x_3 + \\
 &\quad + [\kappa_{23}\kappa_{9,10}\kappa_{89}\kappa_{12}(\kappa_{12,11} + \kappa_{12,13}) + \kappa_{11,12}\kappa_{21}\kappa_{12,13}(\kappa_{98} + \kappa_{9,10})(\kappa_{32} + \kappa_{34})]x_8 \\
 g_4 &= \kappa_{67}x_6 - \kappa_{34}x_3 \\
 g_5 &= \kappa_{56}\kappa_{67}x_4x_5 + \kappa_{34}(-\kappa_{65} - \kappa_{67})x_3 \\
 g_6 &= \kappa_{23}x_2 + (-\kappa_{32} - \kappa_{34})x_3 \\
 g_7 &= -\kappa_{21}(\kappa_{32} + \kappa_{34})x_3 + \kappa_{12}\kappa_{23}x_1
 \end{aligned}$$

Note that  $g_1$  has the form  $g_1 = Q_1x_3 - Q_2x_3x_7 = x_3(Q_1 - Q_2x_7)$ , where  $Q_1$  and  $Q_2$  are homogeneous polynomials in the reaction rate constants which are positive for positive values of the  $\kappa_{ij}$ . Thus, *any* positive steady state satisfies that the value of  $x_7 = x_{Y_p}$  equals  $Q_1/Q_2$ , independently of the initial concentrations. So the Level 2 invariant  $g_1 \in I_f$  shows immediately that the system exhibits *ACR* in  $Y_p$ .

Note that the two monomials that occur in  $g_1$  correspond to two complexes in our network, so one could imagine that it is possible to get a binomial only involving  $x_3x_7$  and  $x_7$  as a Level 1 invariant, that is, via  $\mathbb{R}$ -linear combinations of  $f_1, \dots, f_9$ . However, we prove in [69] that this is not possible and that Level 1 invariants cannot reveal the *ACR* property.

**4.2. Invariants and robust bounds.** Most of the literature on chemical reaction networks only deals with the computation of Level 1 steady state invariants, that we call Type 1 Complex Invariants in [58], from where we extracted the following examples of CRN with different *bifunctional* enzymes. Being bifunctional means that the same enzyme has two

different binding sites in such a way that the enzyme can both catalyze a phosphorylation, or the reverse dephosphorylation.

The first example is a biologically plausible modification of the network (9): we add a reaction with the self dephosphorylation  $Y_p \longrightarrow Y$  of  $Y_p$ . The resulting system does not have toric steady states. We show in [58] that there is no longer ACR behaviour in the variable  $x_7 = x_{Y_p}$ , but instead we found a particular Level 1 invariant depending on a selected proper subset of the complexes, that allowed us to find two nontrivial bounds at steady state independent of initial conditions and the total amounts of the enzymes. According to the numbering of the nodes in [58], we have that for any positive steady state,

$$x_{Y_p} < \min\left\{\frac{k_1}{k_2} \frac{k_3 k_5}{(k_4 + k_5)} \frac{(k_{14} + k_{15})}{k_{13} k_{15}}, k_5 \frac{(k_{11} + k_{12})}{k_{10} k_{12}}\right\}.$$

The factors in these bounds as well as the particular reaction rate constants entering the expressions can be biochemically interpreted.

The second example corresponds to a bifunctional enzyme (known as PFK2-F2,6BPase) in a mammalian cell. In this case, we again get in [58] from Level 1 particular invariants depending on some chosen complexes and the signs of their coefficients, a robust bound in the concentration of a smaller enzyme known as fructose-6-phosphate or in the concentration of the enzyme called F2,6BP, depending on the sign of the difference of two specific reaction rate constants, independently of the stoichiometric compatibility class [58, 18].

**4.3. Invariants and implicit dose–response curves.** We say that  $b > 0$  is a *trivial upper bound* for the  $i$ th species if there exists a conservation relation  $a_1 x_1 + a_2 x_2 + \cdots + x_i + \cdots + a_s x_s = b$  with all  $a_j \geq 0$ . Note that  $b$  is an upper bound for the concentration of  $x_i$  *all along the trajectory*. Using invariants and elimination, we show in [67] how to improve these bounds for steady state concentrations of specific species of the system that are usually considered as the *output*, and so we bound what is called the *maximal response* of the system, regardless of the occurrence of multistationarity. For example, the concentration of the doubly phosphorylated substrate  $x_3 = x_{S_2}$  can be taken as the output in the sequential phosphorylation mechanism with two sites. The *input* of the system is in general a quantity that depends on the initial concentrations, for instance the total amount  $E_{tot}$  of the kinase in (5), that can be usually regulated.

Denote by  $\sigma$  the codimension of the stoichiometric subspace  $S$  and suppose  $f_1, \dots, f_{s-\sigma}$  are linearly independent. Choose also  $\sigma$  independent conservation relations  $\ell_1 - c_1, \dots, \ell_\sigma - c_\sigma$  (where  $\ell_1, \dots, \ell_\sigma$  are homogeneous linear forms and  $c_i$  are constants). Fix the values of  $c_2, \dots, c_\sigma$  and take  $c = c_1$  as our input and  $x_1$  as our output variable. We assume, as it is in general tacitly assumed, that there are a (nonzero) finite number of (complex) solutions to the equations

$$(10) \quad f_1 = f_2 = \cdots = f_\sigma = \ell_1 - c = \ell_2 - c_2 = \cdots = \ell_\sigma - c_\sigma = 0,$$

for any value of  $c$ . In particular, there are a (nonzero) finite number of points in the intersection of  $V(f)$  with each stoichiometric compatibility class. It can be seen that there exists a nonzero polynomial  $p = p(c, x_1)$  in the ideal generated by the polynomials  $f_1, f_2, \dots, f_s, \ell_1 - c, \ell_2 - c_2, \dots, \ell_\sigma - c_\sigma$  in  $\mathbb{R}[c, x_1, \dots, x_s]$ , depending only on  $x_1$  and  $c$  and with positive degree in  $x_1$ , which can be computed with standard elimination tools in computational algebraic geometry (see Lemma 2.1 in [67]). The curve  $\mathcal{C} = \{p = 0\}$  gives

the implicit relation between the input and output variables at steady state, that we call an *implicit dose-response* curve, extending the name of dose-response curve usually given in case  $x_1$  can be analytically expressed in terms of  $c$ . In the general case,  $p$  has high degree both in  $c$  and in  $x_1$  and no such expression is available. However, if one is able to plot the curve  $\mathcal{C} = \{p = 0\}$ , then an upper bound for the values of  $x_1$  at steady state can be read from this plotting, but an implicit plot has in general bad quality and is inaccurate. Instead, one can appeal to the properties of resultants and discriminants to preview a “box” containing the intersection of  $\mathcal{C}$  with the first orthant in the plane  $(x_1, c)$ . This gives improved bounds which yield smaller starting boxes to launch numerical computations. We moreover illustrate in the application to the enzymatic network studied in [62], the relation between the exact implicit dose-response curve we obtain symbolically and the standard hysteresis diagram provided by a numerical solver that is currently seen in the literature. The setting and tools we propose in [67] could yield many other results adapted to any autonomous polynomial dynamical system.

## 5. GENERAL RESULTS ON SIGN CONDITIONS AND MULTISTATIONARITY

Uniqueness of positive solutions plays an important role in many applications and domains of mathematics, beyond chemical reaction networks. In the recent joint paper [65], we were able to isolate and generalize many previous results, in particular, Birch’s theorem [2] in Statistics and Feinberg’s theorem for complex balanced equilibria in case of deficiency zero [27, Prop. 5.3 and Cor. 5.4], as well as Theorem 3.1 above (together with several other results quoted in [65]). The setting is as follows, where  $n$  and  $m$  refer to any two natural numbers (so  $m$  does not denote in this section the number of complexes, and  $n$  could be  $s, d$  or any other suitable number of variables).

Consider a family of generalized polynomial maps  $f_\kappa : \mathbb{R}_{>0}^n \rightarrow \mathbb{R}^m$  defined on the positive orthant, associated with two fixed real matrices  $A = (a_{ij}) \in \mathbb{R}^{m \times r}$ ,  $B = (b_{ij}) \in \mathbb{R}^{r \times n}$ , and  $r$  real positive parameters  $\kappa \in \mathbb{R}_{>0}^r$ :

$$(11) \quad f_{\kappa,i}(x) = \sum_{j=1}^r a_{ij} \kappa_j x_1^{b_{j1}} \dots x_n^{b_{jn}}, \quad i = 1, \dots, m.$$

Note that we allow real exponents and not only nonnegative integer exponents.

**Definition 5.1.** We say that  $f_\kappa : \mathbb{R}_{>0}^n \rightarrow \mathbb{R}^m$  is injective with respect to a subset  $S \subset \mathbb{R}^n$  if for all distinct  $x, y \in \mathbb{R}_{>0}^n$  such that  $x - y \in S$  we have  $f(x) \neq f(y)$ .

Clearly, if we have a mass-action kinetics system (2), which is injective with respect to the stoichiometric subspace  $S$  according to Definition 5.1 with  $f_\kappa = f$ , then there cannot be two different positive steady states on any stoichiometric compatibility class.

The following result is a simplified version of Theorem 1.4 in [65], where we also discuss the algorithmic issues. We need the following notations. The *sign vector*  $\sigma(x) \in \{-, 0, +\}^n$  of a vector  $x \in \mathbb{R}^n$  is defined componentwise. Given a subset  $T$ ,  $\sigma(T)$  denotes the set of sign vectors of all elements in  $T$  and  $\Sigma(T^*) = \sigma^{-1}(\sigma(T) \setminus \{0\})$  denotes the set of all vectors with the same sign of some nonzero vector in  $T$ .

**Theorem 5.2** ([65]). *The following statements are equivalent:*

- (inj) *The map  $f_\kappa$  is injective with respect to  $S$ , for all  $\kappa \in \mathbb{R}_{>0}^r$ .*

(sig)  $\sigma(\ker(A)) \cap \sigma(B(\Sigma(S^*))) = \emptyset$ .

In the particular case  $m = n$  with  $\text{rank}(A)=\text{rank}(B)=n$  and  $S = \mathbb{R}^n$ , condition (sig) simply reads  $\sigma(\ker(A)) \cap \sigma(\text{im}(B)) = \{0\}$ . In oriented matroid language [70], this can be phrased as: no nonzero vector of  $A$  is orthogonal to all vectors of  $B^T$ , or, equivalently, no nonzero covector of  $B^T$  is orthogonal to all covectors of  $A$ . In this framework, we recognized in [65] the first partial version of Descartes’ rule of signs, proposed by René Descartes in 1637 in “La Géométrie”, an appendix to his “Discours de la Méthode”. No multivariate generalization is known and only a lower bound together with a disproven conjecture was proposed in [56].

Recall that Descartes’ rule of signs says that given a univariate real polynomial  $f(x) = a_0 + \sum_{j=1}^r a_j x^j$ , the number of positive real roots of  $f$  is bounded above by the number  $n_f$  of *sign variations* in the ordered sequence of coefficient signs  $\sigma(a_0), \dots, \sigma(a_r)$  (where we discard the 0’s in this sequence and we add a 1 each time two consecutive signs are different). For instance, if  $f = a_0 + 3x - 90x^6 + 2x^8 + x^{111}$ , the sequence of coefficient signs (discarding 0’s) is:  $\sigma(a_0), +, -, +, +$ . So,  $n_f$  equals 2 if  $a_0 \geq 0$  and 3 if  $a_0 < 0$ . Then,  $f$  has at most 2 or 3 positive real roots. This bound is true in case of real, not necessarily natural, exponents. Note that being a condition only depending on the sign of the coefficients, the consequence should also hold for any other polynomial with the same vector of signs, that is, for any polynomial of the form  $f_\kappa(x) = a_0 + \sum_{j=1}^r a_j \kappa_j x^j$ , for any choice of positive  $\kappa \in \mathbb{R}_{>0}^r$ .

The partial multivariate generalization is as follows. Given matrices  $A \in \mathbb{R}^{n \times r}$ ,  $B \in \mathbb{R}^{r \times n}$  with  $n \leq r$  and any index set  $J \subseteq \{1, \dots, r\}$  of cardinality  $n$ , we denote by  $\det(A_J)$  (resp.  $\det(B_J)$ ) the minor indexed by the columns (resp. rows) in  $J$ . For any choice of  $y \in \mathbb{R}^n$ , we consider the system of  $n$  equations in  $n$  unknowns

$$(12) \quad \sum_{j=1}^r a_{ij} x_1^{b_{j1}} \cdots x_n^{b_{jn}} = y_i, \quad i = 1, \dots, n.$$

We denote by  $C^\circ(A)$  the interior of the polyhedral cone generated by the column vectors  $a^1, \dots, a^r$  of  $A$ :

$$C^\circ(A) = \left\{ \sum_{i=1}^r \mu_i a^i \in \mathbb{R}^n : \mu \in \mathbb{R}_+^r \right\}.$$

Item (bnd) in the following result from [65] was previously found (but not highlighted) in [14]. Item (surj) is based on Theorem 3.8 in [64].

**Theorem 5.3.** *[Multivariate Descartes’ rule for (at most) one positive real root] Let  $A \in \mathbb{R}^{n \times r}$  and  $B \in \mathbb{R}^{r \times n}$  be matrices of rank  $n$ . Then,*

- (bnd) *Assume that for all index sets  $J$  of cardinality  $n$ , the product  $\det(A_J) \det(B_J)$  either is zero or has the same sign as all other nonzero products, and moreover, at least one such product is nonzero. Then, (12) has at most one positive solution  $x \in \mathbb{R}_{>0}^n$ , for any  $y \in \mathbb{R}^n$ .*
- (surj) *Assume that the row vectors of  $B$  lie in an open half-space and that the determinants  $\det(A_J)$  and  $\det(B_J)$  have the same sign for all index sets  $J$  of cardinality  $n$ , or the opposite sign in all cases. Then, (12) has exactly one positive solution  $x \in \mathbb{R}_{>0}^n$  if and only if  $y \in C^\circ(A)$ .*

In particular, if the associated oriented matroids of  $A$  and  $B$  are equal, there is at most one positive solution. Note that in case  $n = 1$ , the conditions in Theorem 5.3 read as follows. For  $f = a_0 + \sum_{j=1}^r a_j x^j \in \mathbb{R}[x]$ ,  $A$  is the  $1 \times r$  matrix with entries  $a_1, a_2, \dots, a_r$ ,  $B$  is the  $r \times 1$  matrix with  $b_{j1} = j$  for all  $j = 1, \dots, r$ ,  $c_1 = -a_0$ . The hypotheses of the theorem reduce to asking that  $a_1, \dots, a_r \geq 0$  (or  $\leq 0$ ) and not all 0. So, there is at most one change sign (depending on  $\sigma(a_0)$ ) and so at most one positive root, as in classical Descartes' rule.

## 6. ACKNOWLEDGMENTS

I am grateful to the organization of the First Mathematical Congress of the Americas for the invitation to speak at this wonderful event and to the editors of this Proceedings Volume. I am also grateful to Murad Banaji, Carsten Conradi and Mercedes Pérez Millán for allowing me to use their figures. I am very thankful to my friends Reinhard Laubenbacher and Bernd Sturmfels, who opened for me the window to biological applications from an algebro-geometric background.

## REFERENCES

- [1] Anderson D. (2012), A proof of the Global Attractor Conjecture in the single linkage class case, *SIAM J. Appl. Math.* 71:4, 1487–1508.
- [2] Birch, M.W. (1963), Maximum likelihood in three-way contingency tables, *J. Roy. Stat. Soc. B Met.* 25, 220–233.
- [3] Boulier, F. Software BLAD, <http://www.lifl.fr/~boulier/pmwiki/pmwiki.php/Main/BLAD>.
- [4] Boulier, F. (2007), Differential elimination and biological modelling, in: Gröbner bases in symbolic analysis, *Radon Ser. Comput. Appl. Math.* 2, Walter de Gruyter, 109–137.
- [5] Boulier, F., Lemaire, F., Petitot, M., Sedoglavic, A. (2011) Chemical Reaction Systems, *Computer Algebra and Systems Biology*, in: *Proceedings of Computer Algebra in Scientific Computing, LNCS 6885*, Vladimir Gerdt et al. (eds.), <http://hal.archives-ouvertes.fr/hal-00603290>, 73–87.
- [6] Boulier, F., Lefranc, M., Lemaire, F. and Morant, P.-E. (2011), Model reduction of chemical reaction systems using elimination, *Math. Comput. Sci.* 5(3), 289–301.
- [7] Lemaire, F., Ürgüplü, A. (2012), MABSys: modeling and analysis of biological systems, in: *Algebraic and numeric biology, Lecture Notes in Comput. Sci.* 6479, Springer, 57–75.
- [8] Bykov V. , Kytmanov A., Lazman M., with the edition of Passare M. (1998), *Elimination Methods in Polynomial Computer Algebra Mathematics and Its Applications*, vol. 448 Springer.
- [9] Conradi C, Flockerzi D, Raisch J, Stelling J. (2007), Subnetwork analysis reveals dynamic features of complex (bio)chemical networks, *Proc Natl Acad Sci U S A.* 104(49), 19175–19180. Parametrizing the relevant region in parameter space,
- [10] Cox D., Little J., O'Shea D. (2007), *Ideals, Varieties and Algorithms*, 3rd. edition, Springer.
- [11] Craciun G., (2015), Toric Differential Inclusions and a Proof of the Global Attractor Conjecture. *ArXiv:1501.02860*.
- [12] Craciun G., Dickenstein A., Shiu A., Sturmfels B. (2009), Toric dynamical systems, *Journal of Symbolic Computation* 44, 1551–1565.
- [13] Craciun G., Feinberg M. (2006), Multiple Equilibria in Complex Chemical Reaction Networks: II. The Species-Reactions Graph, *SIAM Journal on Applied Mathematics* 66(4), 1321–1338.
- [14] Craciun G., Garcia-Puente L., Sottile F. (2010), Some Geometric Aspects of Control Points for Toric Patches, *Proceedings of the 7th international conference on Mathematical Methods for Curves and Surfaces*, (M. Daehlen et al. Eds), *Lecture Notes in Computer Science* 5862, Springer, 111–135.
- [15] Craciun G., Tang Y., Feinberg M. (2006), Understanding bistability in complex enzymedriven reaction networks, *Proc. Natl. Acad. Sci.* 103(23), 8697–8702.
- [16] Craciun G., Nazarov F., Pantea C. (2013), Persistence and permanence of mass-action and power-law dynamical systems, *SIAM Journal on Applied Mathematics* 73:1.



- [17] Craciun G., Pantea C. (2008), Identifiability of chemical reaction networks. *J. Math. Chem.* 44, 244–259.
- [18] Dasgupta T., Croll D. H., Owen J.A., Vander Heiden M. G., Locasale J. W., Alon U., Cantley L. C., Gunawardena J. (2013), A fundamental trade off in covalent switching and its circumvention in glucose homeostasis, Manuscript.
- [19] Dickenstein A., Feliu, E. (2014), Algebraic methods for biochemical reaction networks, Book in preparation.
- [20] Decker W., Greuel G.-M., Pfister G., Schönemann H. (2011), SINGULAR 3-1-3 — A computer algebra system for polynomial computations. <http://www.singular.uni-kl.de>.
- [21] Deng J., Jones C., Feinberg M., Nachman A. (2011), On the steady states of weakly reversible chemical reaction networks, Available at arXiv:1111.2386.
- [22] Dickenstein A., Pérez Millán M. (2011), How far is complex balancing from detailed balancing? *Bulletin of Mathematical Biology* 73(4), 811–828.
- [23] Domijian, M., Kirkilionis, M. (2009), Bistability and oscillations in chemical reaction networks, *J. Math. Biol.* 58(4), 467–501.
- [24] Feinberg M. (1972), Complex balancing in general kinetic systems, *Arch. Rational Mech. Anal.* 49(3), 187–194.
- [25] Feinberg M., Horn F. (1977), Chemical mechanism structure and the coincidence of the stoichiometric and kinetic subspaces. *Arch. Rational Mech. Anal.* 66, 83–97.
- [26] Feinberg M. (1977), Mathematical Aspects of Mass Action Kinetics, Chapter 1 in *Chemical Reactor Theory: A Review* (eds. N. Amundson and L. Lapidus), Prentice-Hall.
- [27] Feinberg M. (1979), Lectures on chemical reaction networks, U. of Wisconsin, available at: <http://www.che.eng.ohio-state.edu/~FEINBERG/LecturesOnReactionNetworks>.
- [28] Feinberg M. (1989), Necessary and sufficient conditions for detailed balancing in mass action systems of arbitrary complexity, *Chem. Eng. Sci.* 44(9), 1819–1827.
- [29] Feinberg M. (1995), The existence and uniqueness of steady states for a class of chemical reaction networks, *Arch. Rational Mech. Anal.* 132(4), 311–370.
- [30] Feinberg M. (1995), Multiple steady states for chemical reaction networks of deficiency one, *Arch. Rational Mech. Anal.* 132(4), 371–406.
- [31] Feliu E., Knudsen M., Andersen L.N., Wiuf C. (2012), An Algebraic Approach to Signaling Cascades with  $n$  Layers, *B. Math. Biol.* 74(1), 45–72.
- [32] Feliu E., Wiuf C. (2012), Variable elimination in chemical reaction networks with mass action kinetics, *SIAM Journal on Applied Mathematics* 72:4, 959–981.
- [33] Feliu E., Wiuf C. (2012), Enzyme sharing as a cause of multistationarity in signaling systems, *Journal of the Royal Society Interface* 9:71, 1224–32.
- [34] Feliu E., Wiuf C. (2013), Variable elimination in post-translational modification reaction networks with mass-action kinetics, *Journal of Mathematical Biology* 66:1, 281–310.
- [35] Feliu E., Wiuf C. (2013), Simplifying Biochemical Models With Intermediate Species, *Journal of the Royal Society Interface*, 10:20130484.
- [36] Ferrell, J.E. Jr. (2009), Q&A: Systems biology, *J. Biol.* 8(1), art. 2.
- [37] Flockerzi, D. Holstein, K. Conradi, C. (2014),  $N$ -site phosphorylation systems with  $2N-1$  steady states, *Bull. Math. Biol.* 76(8), 1892–1916.
- [38] Gatermann K. (2001), Counting stable solutions of sparse polynomial systems in chemistry, *Contemporary Mathematics*, Volume 286, Symbolic Computation: Solving Equations in Algebra, Geometry and Engineering, (Editors E. Green et al.), 53–69.
- [39] Gatermann K., Huber B. (2002), A family of sparse polynomial systems arising in chemical reaction systems, *J. Symbolic Comput.* 33(3), 275–305.
- [40] Gatermann K., Wolfrum M. (2005), Bernstein’s second theorem and Viro’s method for sparse polynomial systems in chemistry, *Advances in Applied Mathematics* 34(2):252–294.
- [41] Grayson D. R., Stillman M. E. Macaulay2, a software system for research in algebraic geometry. Available at <http://www.math.uiuc.edu/Macaulay2/>
- [42] Gnacadja G. (2009), Univalent Positive Polynomial Maps and the Equilibrium State of Chemical Networks of Reversible Binding Reactions, *Adv. Appl. Math.* 43(4), 394–414.
- [43] Gopalkrishnan M. (2011), Catalysis in Reaction Network, *Bull. Math. Biol.* 73, 2962–2982.

- [44] Gorban A., Yablonsky G. (2011) Extended detailed balance for systems with irreversible reactions, *Chemical Engineering Science* 66, 5388–5399.
- [45] Gunawardena J. (2003), Chemical reaction network theory for in-silico biologists, Technical Report, Available at: <http://vcp.med.harvard.edu/papers/crnt.pdf>.
- [46] Gunawardena J. (2007), Distributivity and processivity in multisite phosphorylation can be distinguished through steady-state invariants, *Biophys. J.* 93, 3828–3834.
- [47] Gunawardena, J. (2012), A linear framework for time-scale separation in nonlinear biochemical systems, *PLoS ONE* 7, e36321.
- [48] Harrington H., Feliu E. Wiuf C. Stumpf M.P.H. (2013) , Cellular compartments cause multistability in biochemical reaction networks and allow cells to process more information, *Biophysical Journal* 104:8, 1824–1831.
- [49] Hárs V., Tóth J. (1979), On the inverse problem of reaction kinetics, *Colloquia Mathematica Societatis János Bolyai* 30. Qualitative Theory of Differential Equations. Szeged (Hungary), 363–379.
- [50] Horn F. (1972), Necessary and sufficient conditions for complex balancing in chemical kinetics, *Arch. Ration. Mech. Anal.* 49(3), 172–186.
- [51] Horn F., Jackson R. (1972), General mass action kinetics, *Arch. Ration. Mech. Anal.* 47(2), 81–116.
- [52] Horn F. (1973), Stability and complex balancing in mas-action systems with three short complexes, *Proc. Roy. Soc. (London) Ser. A* 334, 331–342.
- [53] Horn F. (1974), in: *Mathematical aspects of chemical and biochemical problems and quantum chemistry* (Proc. SIAM-AMS Sympos. Appl. Math. New York, 1974), SIAM-AMS Proc. Vol. VIII. Amer. Math. Soc. Providence, RI, 125–137.
- [54] Johnston M., Siegel D. (2011), Linear Conjugacy of Chemical Reaction Networks, *J. Math. Chem.* 49(7), 1263–1282.
- [55] Gopalkrishnan M., Miller E., Shiu A. (2014), A geometric approach to the global attractor conjecture, *SIAM J. Appl. Dyn. Syst.* 13 (2), 758–797.
- [56] Itenberg I., Roy M.F. (1996), Multivariate Descartes' rule, *Beiträge zur Algebra und Geometrie* 37(2), 337–346.
- [57] Joshi B. Shiu A. (2013), Atoms of multistationarity in chemical reaction networks, *Journal of Mathematical Chemistry* 51:1, 153–178.
- [58] Karp R., Pérez Millán M., Dasgupta T., Dickenstein A., Gunawardena J. (2012), Complex-linear invariants of biochemical networks, *J. Theor. Biol.* 211, 130–138.
- [59] Kaufman M., Soulé C., Thomas R. (2007), A new necessary condition on interaction graphs for multistationarity, *J. Theor. Biol.* 248(4), 675–85.
- [60] Knudsen M., Feliu E., Wiuf C. (2012), Exact Analysis of Intrinsic Qualitative Features of Phosphorelays using Mathematical Models, *J. Theor. Biology* 300, 7–18.
- [61] Manrai A. K. Gunawardena J. (2008), The geometry of multisite phosphorylation *Biophys. J.* 95(12), 5533–5543.
- [62] Markevich N. I., Hoek J. B., Kholodenko B. N. (2004), Signaling switches and bistability arising from multisite phosphorylation in protein kinase cascades, *J. Cell Biol.* 164(3), 353–359.
- [63] McKeithan, T. (1995), Kinetic proofreading in T-cell receptor signal transduction, *Proc. Natl. Acad. Sci. USA* 92(11), 5042–5046.
- [64] Müller S., Regensburger G. (2012), Generalized mass action systems: Complex balancing equilibria and sign vectors of the stoichiometric and kinetic-order subspaces, *SIAM J. Appl. Math.* 72(6), 1926–1947.
- [65] Müller S., Feliu E., Regensburger G., Conradi C., Shiu A., Dickenstein A. (2015), Sign conditions for injectivity of generalized polynomial maps with applications to chemical reaction networks and real algebraic geometry, To appear: *Found. Comput. Math.*
- [66] Pantea C., Gupta A., Rawlings J., Craciun G. (2013), The QSSA in chemical kinetics: As taught and as practiced, in: Jonoska, N and Saito, M (Eds.), *Discrete and Topological Models in Molecular Biology*, Springer.
- [67] Pérez Millán M., Dickenstein A. (2014), Implicit dose-response curves, To appear: *J. Math. Biol.*
- [68] Pérez Millán M., Dickenstein A. (2014), Steady states of MESSI biological systems, Manuscript.
- [69] Pérez Millán M., Dickenstein A., Shiu A., Conradi C. (2012), Chemical reaction systems with toric steady states, *Bulletin of Mathematical Biology* 74(5), 1027–1065.

- [70] Richter-Gebert J., Ziegler, G. (1997), Oriented matroids, in: Handbook of discrete and computational geometry, CRC Press, 111–132.
- [71] Shen-Orr S.S., Milo R., Mangan S., Alon U. (2002), Network motifs in the transcriptional regulation network of *Escherichia coli*, *Nat Genet.* 31(1), 64–68.
- [72] Shinar G. Milo R. Martínez M. R. Alon U. (2007), Input-output robustness in simple bacterial signaling systems. *Proc. Natl. Acad. Sci. USA* 104, 19931–19935.
- [73] Shinar G., Alon U., Feinberg M. (2009), Sensitivity and robustness in chemical reaction networks, *Siam J. Appl. Math.* 69(4), 977–998.
- [74] Shinar G., Feinberg M. (2010), Structural sources of robustness in biochemical reaction networks, *Science* 327(5971), 1389–1391.
- [75] Shiu A., Sturmfels B. (2010), Siphons in chemical reaction networks, *Bulletin of Mathematical Biology* 72: 1448–1463.
- [76] Sontag E. (2001), Structure and stability of certain chemical networks and applications to the kinetic proofreading model of T-cell receptor signal transduction, *IEEE Trans. Automat. Control* 46:1028–1047.
- [77] Thomson M., Gunawardena J. (2009), The rational parameterisation theorem for multisite post-translational modification systems, *J. Theor. Biol.* 261(4), 626–636.
- [78] Thomson M., Gunawardena J. (2009), Unlimited multistability in multisite phosphorylation systems, *Nature* 460(7252), 274–277.
- [79] Vol’pert A., Hudjaev S. (1985), Analysis in classes of discontinuous functions and equations of mathematical physics, volume 8 of *Mechanics: Analysis*, Martinus Nijhoff Publishers.
- [80] Wang L., Sontag E. (2008), On the number of steady states in a multiple futile cycle, *J. Math. Biol.* 57(1), 29–52.

Dto. de Matemática, FCEN - Universidad de Buenos Aires, and IMAS (UBA-CONICET), Ciudad Universitaria - Pab. I - C1428EGA Buenos Aires, Argentina

*E-mail address:* `alidick@dm.uba.ar`

*URL:* `http://mate.dm.uba.ar/~alidick`