Families of polynomials in the study of biochemical reaction networks

Alicia Dickenstein

Departamento de Matemática, FCEN,
Universidad de Buenos Aires,
and Instituto de Matemática Luis A. Santaló, UBA-CONICET

MEGA 2021, June 10 2021

Setting for this talk

- (Bio)chemical reaction networks define systems of ordinary differential equations with (in general unknown) parameters
- The basic mathematical theory was developed by chemical engineers: Horn, Jackson y Feinberg, since 1972. Tools from (real and complex) algebraic geometry are being used since the pioneering work of Karin Gatermann '01-'04.
- The standard assumption is mass-action kinetics (when there are enough molecules and well mixed...). In this case, we get families of polynomial ordinary differential equations, with a combinatorial structure that comes from the digraph of reactions.

Setting for this talk

- (Bio)chemical reaction networks define systems of ordinary differential equations with (in general unknown) parameters
- The basic mathematical theory was developed by chemical engineers: Horn, Jackson y Feinberg, since 1972. Tools from (real and complex) algebraic geometry are being used since the pioneering work of Karin Gatermann '01-'04.
- The standard assumption is mass-action kinetics (when there are enough molecules and well mixed...). In this case, we get families of polynomial ordinary differential equations, with a combinatorial structure that comes from the digraph of reactions.

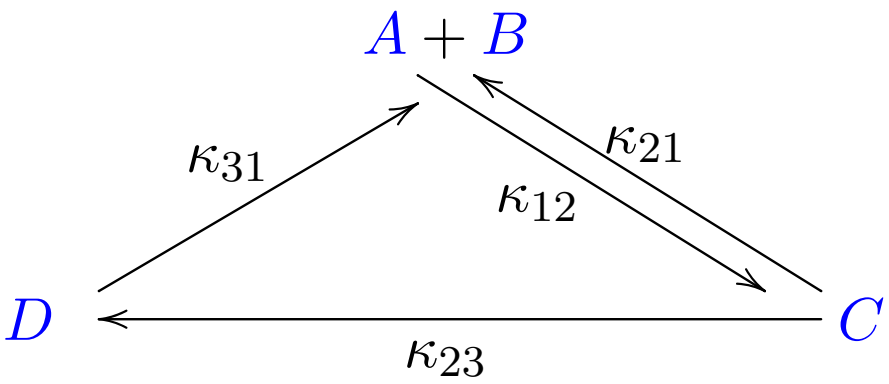
Setting for this talk

- (Bio)chemical reaction networks define systems of ordinary differential equations with (in general unknown) parameters
- The basic mathematical theory was developed by chemical engineers: Horn, Jackson y Feinberg, since 1972. Tools from (real and complex) algebraic geometry are being used since the pioneering work of Karin Gatermann '01-'04.
- The standard assumption is mass-action kinetics (when there are enough molecules and well mixed...). In this case, we get families of polynomial ordinary differential equations, with a combinatorial structure that comes from the digraph of reactions.

Example: T-cell signal transduction model

T-cell receptors bind to both self-antigens and foreign antigens. *How can T-cells be sensitive and specific in recognizing self vs. foreign?*

Model due to [McKeithan '95], immunologist; [Sontag '01]:



- A = T-cell receptor,
  B = MHC of antigen-presenting cell
- C = A bound to B,
  D = activated form of C

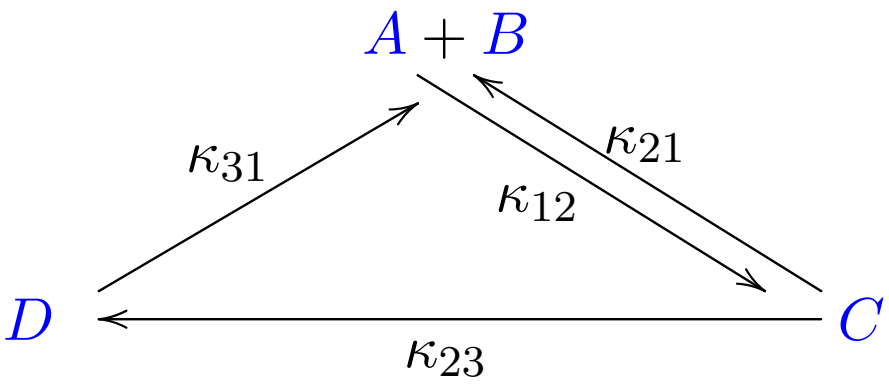
This *CRN* has:

- 4 *reactions* among the...
- $m = 3$ **complexes** $A + B$, C, and D which are composed by...
- $s = 4$ **species** A, B, C, and D.

Example: T-cell signal transduction model

T-cell receptors bind to both self-antigens and foreign antigens. *How can T-cells be sensitive and specific in recognizing self vs. foreign?*

Model due to [McKeithan '95], immunologist; [Sontag '01]:

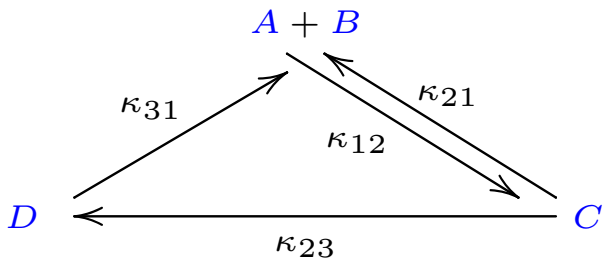


- A = T-cell receptor,
  B = MHC of antigen-presenting cell
- C = A bound to B,
  D = activated form of C

This *CRN* has:

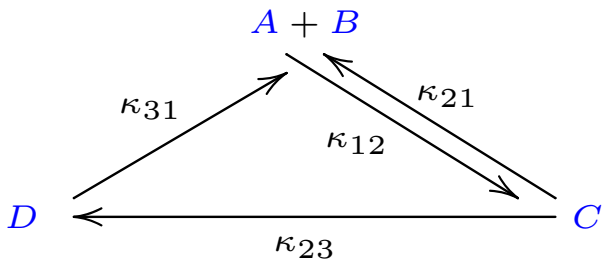
- 4 *reactions* among the...
- $m = 3$ **complexes** $A + B$, C, and D which are composed by...
- $s = 4$ **species** A, B, C, and D.

Explicit equations governing the concentrations $x(t) = (x_A(t), x_B(t), x_C(t), x_D(t))$



$$\frac{dx}{dt} = \kappa_{12}\, x_A x_B \begin{pmatrix} -1 \\ -1 \\ 1 \\ 0 \end{pmatrix} + \kappa_{21}\, x_C \begin{pmatrix} 1 \\ 1 \\ -1 \\ 0 \end{pmatrix} + \kappa_{23}\, x_C \begin{pmatrix} 0 \\ 0 \\ -1 \\ 0 \end{pmatrix} + \kappa_{31} x_D \begin{pmatrix} 1 \\ 1 \\ 0 \\ -1 \end{pmatrix}$$

$$\begin{aligned}
\frac{dx_A}{dt} &= f_1(x) = -\kappa_{12} x_A x_B + \kappa_{21} x_C + \kappa_{31} x_D \\
\frac{dx_B}{dt} &= f_2(x) = -\kappa_{12} x_A x_B + \kappa_{21} x_C + \kappa_{31} x_D \\
\frac{dx_C}{dt} &= f_3(x) = \kappa_{12} x_A x_B - \kappa_{21} x_C - \kappa_{23} x_C \\
\frac{dx_D}{dt} &= f_4(x) = \kappa_{23} x_C - \kappa_{31} x_D
\end{aligned}$$



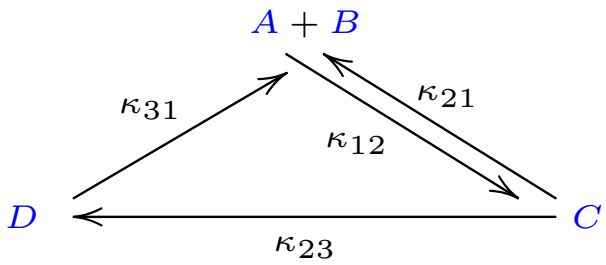
$$
\begin{aligned}
\frac{dx_A}{dt} &= f_1(x) = -\kappa_{12} x_A x_B + \kappa_{21} x_C + \kappa_{31} x_D \\
\frac{dx_B}{dt} &= f_2(x) = -\kappa_{12} x_A x_B + \kappa_{21} x_C + \kappa_{31} x_D \\
\frac{dx_C}{dt} &= f_3(x) = \kappa_{12} x_A x_B - \kappa_{21} x_C - \kappa_{23} x_C \\
\frac{dx_D}{dt} &= f_4(x) = \kappa_{23} x_C - \kappa_{31} x_D
\end{aligned}
$$

Linear dependencies give conservation relations

From $f_1 + f_3 + f_4 = f_2 + f_3 + f_4 = 0$, we get two conservation relations:

$$
\begin{aligned}
x_A + x_C + x_D &= T_1, \\
x_B + x_C + x_D &= T_2.
\end{aligned}
$$

Thus, trajectories lie in a 2-plane in 4-space. Total amounts T_1, T_2 are determined by the initial conditions $x(0)$.



$$\frac{dx_A}{dt} = f_1(x) = -\kappa_{12}x_Ax_B + \kappa_{21}x_C + \kappa_{31}x_D$$
$$\frac{dx_B}{dt} = f_2(x) = -\kappa_{12}x_Ax_B + \kappa_{21}x_C + \kappa_{31}x_D$$
$$\frac{dx_C}{dt} = f_3(x) = \kappa_{12}x_Ax_B - \kappa_{21}x_C - \kappa_{23}x_C$$
$$\frac{dx_D}{dt} = f_4(x) = \kappa_{23}x_C - \kappa_{31}x_D$$

Linear dependencies give conservation relations

From $f_1 + f_3 + f_4 = f_2 + f_3 + f_4 = 0$, we get two conservation relations:

$$x_A + x_C + x_D = T_1,$$
$$x_B + x_C + x_D = T_2.$$

Thus, trajectories lie in a 2-plane in 4-space. Total amounts T_1, T_2 are determined by the initial conditions $x(0)$.

CRN with mass-action kinetics

- Starting data: a set of r reactions (labeled edges, e.g. $i \overset{\kappa_{ij}}{\to} j$, where $\kappa_{ij} \in \mathbb{R}_{>0}$ are the reaction rate constants) between m complexes (giving rise to monomials e.g. $x^{y_i} = x_1^{y_{i1}} x_2^{y_{i2}} \cdots x_s^{y_{is}}$) composed of s species (variables $x_1, \ldots, x_s$).
- A *chemical reaction network* is a finite directed graph $G = (V, E, (\kappa_{ij})_{(i,j)\in E}, (y_i)_{i=1,\ldots,m})$ whose vertices are labeled by complexes and whose edges are labeled by parameters.
- View the concentrations $x_1, x_2, \ldots, x_s$ as *functions of time*.
- *Mass-action kinetics* specified by the network G is the following autonomous system of ordinary differential equations:

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

CRN with mass-action kinetics

- Starting data: a set of r reactions (labeled edges, e.g. $i \overset{\kappa_{ij}}{\to} j$, where $\kappa_{ij} \in \mathbb{R}_{>0}$ are the reaction rate constants) between m complexes (giving rise to monomials e.g. $x^{y_i} = x_1^{y_{i1}} x_2^{y_{i2}} \cdots x_s^{y_{is}}$) composed of s species (variables $x_1, \ldots, x_s$).
- A *chemical reaction network* is a finite directed graph $G = (V, E, (\kappa_{ij})_{(i,j)\in E}, (y_i)_{i=1,\ldots,m})$ whose vertices are labeled by complexes and whose edges are labeled by parameters.
- View the concentrations $x_1, x_2, \ldots, x_s$ as *functions of time*.
- *Mass-action kinetics* specified by the network G is the following autonomous system of ordinary differential equations:

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

CRN with mass-action kinetics

- Starting data: a set of r reactions (labeled edges, e.g. $i \overset{\kappa_{ij}}{\to} j$, where $\kappa_{ij} \in \mathbb{R}_{>0}$ are the reaction rate constants) between m complexes (giving rise to monomials e.g. $x^{y_i} = x_1^{y_{i1}} x_2^{y_{i2}} \cdots x_s^{y_{is}}$) composed of s species (variables $x_1, \ldots, x_s$).
- A *chemical reaction network* is a finite directed graph $G = (V, E, (\kappa_{ij})_{(i,j)\in E}, (y_i)_{i=1,\ldots,m})$ whose vertices are labeled by complexes and whose edges are labeled by parameters.
- View the concentrations $x_1, x_2, \ldots, x_s$ as *functions of time.*
- *Mass-action kinetics* specified by the network G is the following autonomous system of ordinary differential equations:

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

CRN with mass-action kinetics

- Starting data: a set of r reactions (labeled edges, e.g. $i \overset{\kappa_{ij}}{\to} j$, where $\kappa_{ij} \in \mathbb{R}_{>0}$ are the reaction rate constants) between m complexes (giving rise to monomials e.g. $x^{y_i} = x_1^{y_{i1}} x_2^{y_{i2}} \cdots x_s^{y_{is}}$) composed of s species (variables $x_1, \ldots, x_s$).
- A *chemical reaction network* is a finite directed graph $G = (V, E, (\kappa_{ij})_{(i,j)\in E}, (y_i)_{i=1,\ldots,m})$ whose vertices are labeled by complexes and whose edges are labeled by parameters.
- View the concentrations $x_1, x_2, \ldots, x_s$ as *functions of time*.
- *Mass-action kinetics* specified by the network G is the following autonomous system of ordinary differential equations:

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

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

Hungarian Lemma - V. Hárs, J. Tóth, 1979: A polynomial system of s real polynomials $f_1, \dots, f_s$ in s variables arises from a mass-action kinetics dynamical system 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.

An important consequence is that $\mathbb{R}^s_{>0}$ and $\mathbb{R}^s_{\geq 0}$ are forward-invariant for the dynamics

"Chaotic" Lorenz equations are not of this shape, but many models in population dynamics (as the Lotka-Volterra model or the standard epidemiological models) are. CRN in chemistry usually have complexes with high coordinates. Usual models in systems biology, in particular enzymatic pathways, are of this form, with small coordinates.

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

Hungarian Lemma - V. Hárs, J. Tóth, 1979: A polynomial system of s real polynomials $f_1, \dots, f_s$ in s variables arises from a mass-action kinetics dynamical system 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.

An important consequence is that $\mathbb{R}^s_{>0}$ and $\mathbb{R}^s_{\geq 0}$ are forward-invariant for the dynamics

.

"Chaotic" Lorenz equations are not of this shape, but many models in population dynamics (as the Lotka-Volterra model or the standard epidemiological models) are. CRN in chemistry usually have complexes with high coordinates. Usual models in systems biology, in particular enzymatic pathways, are of this form, with small coordinates.

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

Hungarian Lemma - V. Hárs, J. Tóth, 1979: A polynomial system of s real polynomials $f_1, \ldots, f_s$ in s variables arises from a mass-action kinetics dynamical system if and only if there exists real polynomials $p_k, q_k, k = 1, \ldots, s$ with non negative coefficients such that $f_k = p_k - x_k q_k$ for all k.

An important consequence is that $\mathbb{R}^s_{>0}$ and $\mathbb{R}^s_{\geq 0}$ are forward-invariant for the dynamics

.

"Chaotic" Lorenz equations are not of this shape, but many models in population dynamics (as the Lotka-Volterra model or the standard epidemiological models) are. CRN in chemistry usually have complexes with high coordinates. Usual models in systems biology, in particular enzymatic pathways, are of this form, with small coordinates.

Setting for this talk, revised

- We will concentrate on biochemical reaction networks under mass-action kinetics. They give systems of polynomial ordinary differential equations with parameters (κ, T), where κ are reaction rate constants and T are total amounts.
- The reaction rate constants κ are in general unknown or difficult to measure. Standard methods in other sciences involve exhaustive sampling. We look at them as special families of polynomial ordinary differential equations.
- Our aim is to explore the parameter space in order to to predict properties of the associated systems, which usually have too many variables and too many parameters for the standard current computational tools.
- There are many useful mathematical and computational tools, but we are forced to extend the mathematical results and to understand the structure of the networks to make the computations feasible, even for families with an unbounded number of parameters.

Setting for this talk, revised

- We will concentrate on biochemical reaction networks under mass-action kinetics. They give systems of polynomial ordinary differential equations with parameters (κ, T), where κ are reaction rate constants and T are total amounts.
- The reaction rate constants κ are in general unknown or difficult to measure. Standard methods in other sciences involve exhaustive sampling. We look at them as special families of polynomial ordinary differential equations.
- Our aim is to explore the parameter space in order to to predict properties of the associated systems, which usually have too many variables and too many parameters for the standard current computational tools.
- There are many useful mathematical and computational tools, but we are forced to extend the mathematical results and to understand the structure of the networks to make the computations feasible, even for families with an unbounded number of parameters.

Setting for this talk, revised

- We will concentrate on biochemical reaction networks under mass-action kinetics. They give systems of polynomial ordinary differential equations with parameters (κ, T), where κ are reaction rate constants and T are total amounts.
- The reaction rate constants κ are in general unknown or difficult to measure. Standard methods in other sciences involve exhaustive sampling. We look at them as special families of polynomial ordinary differential equations.
- Our aim is to explore the parameter space in order to to predict properties of the associated systems, which usually have too many variables and too many parameters for the standard current computational tools.
- There are many useful mathematical and computational tools, but we are forced to extend the mathematical results and to understand the structure of the networks to make the computations feasible, even for families with an unbounded number of parameters.

Setting for this talk, revised

- We will concentrate on biochemical reaction networks under mass-action kinetics. They give systems of polynomial ordinary differential equations with parameters (κ, T), where κ are reaction rate constants and T are total amounts.
- The reaction rate constants κ are in general unknown or difficult to measure. Standard methods in other sciences involve exhaustive sampling. We look at them as special families of polynomial ordinary differential equations.
- Our aim is to explore the parameter space in order to to predict properties of the associated systems, which usually have too many variables and too many parameters for the standard current computational tools.
- There are many useful mathematical and computational tools, but we are forced to extend the mathematical results and to understand the structure of the networks to make the computations feasible, even for families with an unbounded number of parameters.

Setting for this talk, revised

- We will concentrate on biochemical reaction networks under mass-action kinetics. They give systems of polynomial ordinary differential equations with parameters (κ, T), where κ are reaction rate constants and T are total amounts.
- The reaction rate constants κ are in general unknown or difficult to measure. Standard methods in other sciences involve exhaustive sampling. We look at them as special families of polynomial ordinary differential equations.
- Our aim is to explore the parameter space in order to to predict properties of the associated systems, which usually have too many variables and too many parameters for the standard current computational tools.
- There are many useful mathematical and computational tools, but we are forced to extend the mathematical results and to understand the structure of the networks to make the computations feasible, even for families with an unbounded number of parameters.

Phosho-Dephosphorylation: "futile" cycle

$$S_0 + E \underset{k_{\text{off}}}{\overset{k_{\text{on}}}{\rightleftarrows}} ES_0 \overset{k_{\text{cat}}}{\to} S_1 + E$$

$$S_1 + F \underset{\ell_{\text{off}}}{\overset{\ell_{\text{on}}}{\rightleftarrows}} FS_1 \overset{\ell_{\text{cat}}}{\to} S_0 + F$$

E and F *enzymes*, S_0 and S_1 *substrates*, S_0E and S_1F *intermediates*

and we represent it with: $S_0 \underset{F}{\overset{E}{\rightleftarrows}} S_1$.

There are 6 species, 6 complexes (nodes) and 6 reactions (edges)

The Nobel Prize in Physiology or Medicine 1992 was awarded to Fischer and Krebs "for their discoveries concerning reversible protein phosphorylation as a biological regulatory mechanism".

Phosho-Dephosphorylation: "futile" cycle

$$S_0 + E \underset{k_{\text{off}}}{\overset{k_{\text{on}}}{\rightleftarrows}} ES_0 \overset{k_{\text{cat}}}{\to} S_1 + E$$

$$S_1 + F \underset{\ell_{\text{off}}}{\overset{\ell_{\text{on}}}{\rightleftarrows}} FS_1 \overset{\ell_{\text{cat}}}{\to} S_0 + F$$

E and F *enzymes*, S_0 and S_1 *substrates*, S_0E and S_1F *intermediates*

and we represent it with: $S_0 \underset{F}{\overset{E}{\rightleftarrows}} S_1$.

There are 6 species, 6 complexes (nodes) and 6 reactions (edges)

The Nobel Prize in Physiology or Medicine 1992 was awarded to Fischer and Krebs "for their discoveries concerning reversible protein phosphorylation as a biological regulatory mechanism".

Phosho-Dephosphorylation: "futile" cycle

$$S_0 + E \underset{k_{\text{off}}}{\overset{k_{\text{on}}}{\rightleftarrows}} ES_0 \overset{k_{\text{cat}}}{\to} S_1 + E$$

$$S_1 + F \underset{\ell_{\text{off}}}{\overset{\ell_{\text{on}}}{\rightleftarrows}} FS_1 \overset{\ell_{\text{cat}}}{\to} S_0 + F$$

E and F *enzymes*, S_0 and S_1 *substrates*, S_0E and S_1F *intermediates*

and we represent it with: $S_0 \underset{F}{\overset{E}{\rightleftarrows}} S_1$.

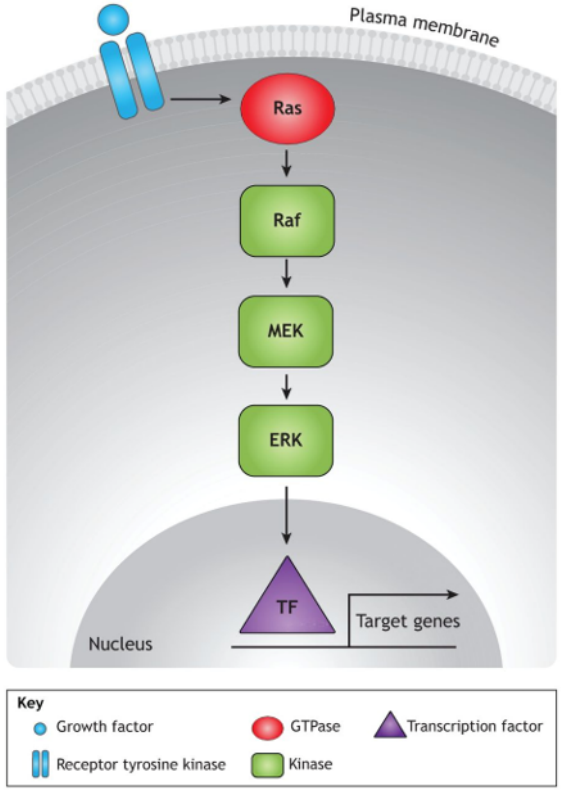
There are 6 species, 6 complexes (nodes) and 6 reactions (edges)

The Nobel Prize in Physiology or Medicine 1992 was awarded to Fischer and Krebs "for their discoveries concerning reversible protein phosphorylation as a biological regulatory mechanism".

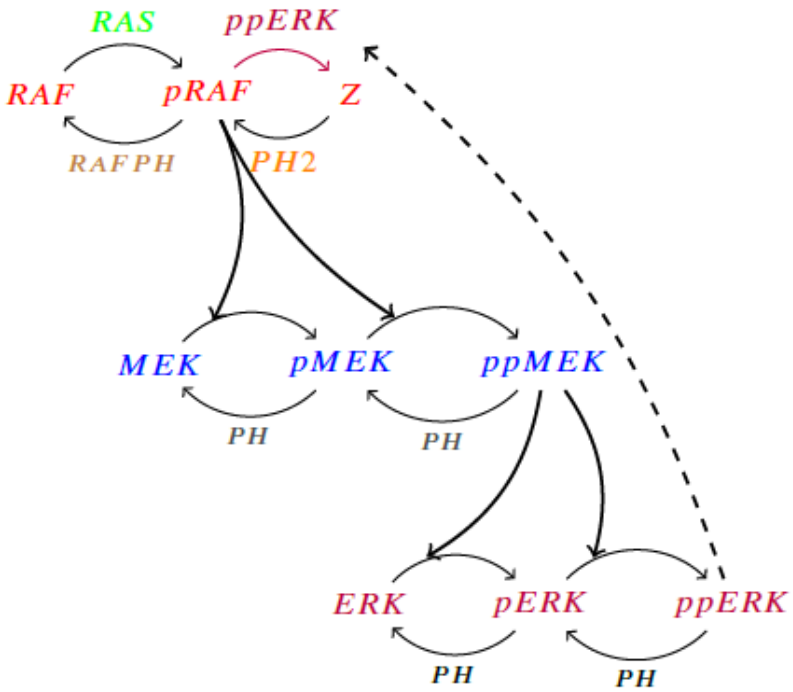
The ERK pathway

"The extracellular signal-regulated kinase (ERK) pathway leads to activation of the effector molecule ERK, which controls downstream responses by phosphorylating a variety of substrates, including transcription factors"

Outstanding questions in developmental ERK Signaling, A. Patel and S. Shvartsman Development (2018) 145, dev143818

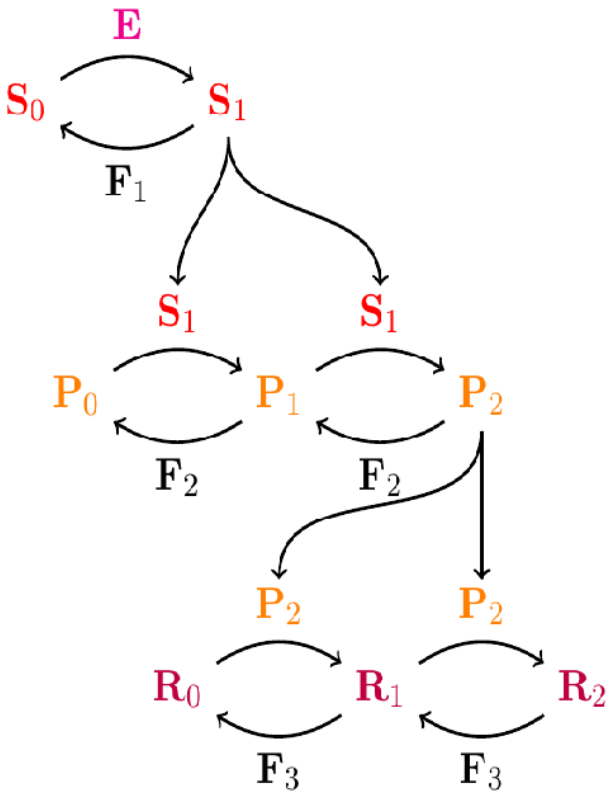


ERK phosphorylation cascade (with or without retroactivity)



The Ras-Raf-MEK-ERK pathway is a cascade of phosphorylation of proteins in the cell that communicates a signal from a receptor on the surface of the cell to the DNA in the nucleus of the cell.
In a cascade, the phosphorylated substrate in one layer acts as an enzyme in the next layer.

ERK phosphorylation cascade without retroactivity



There are 2 +33=8 substrates,

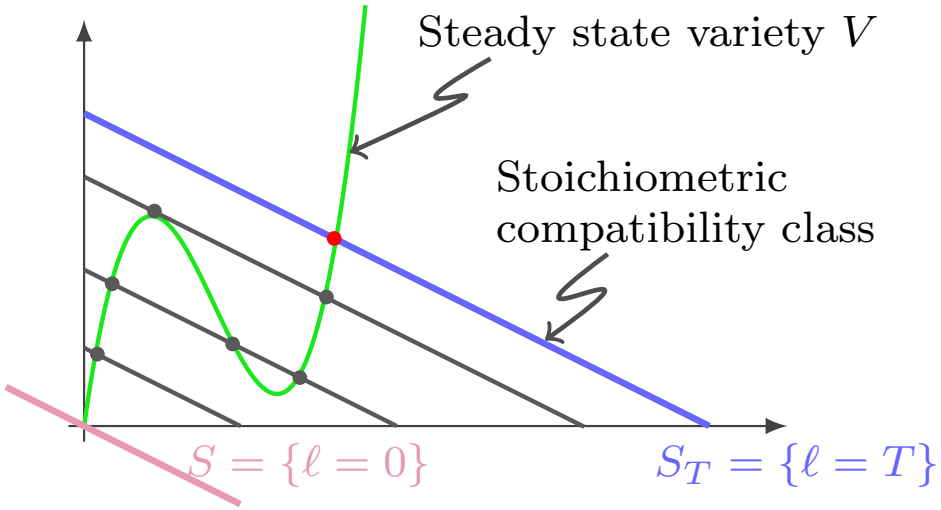
3 or 4 enzymes, 10 intermediates, so 21 or 22 variables. Plus 30 reaction rate constants and 6 or 7 total amounts, so 36 or 37 parameters, depending on whether $F_2 = F_3$ or not.

Then..

How can we study the associated family of polynomial dynamical systems?

Definition

x^* is a steady state of $\frac{dx}{dt} = f(x)$ if $f(x^*) = 0$.

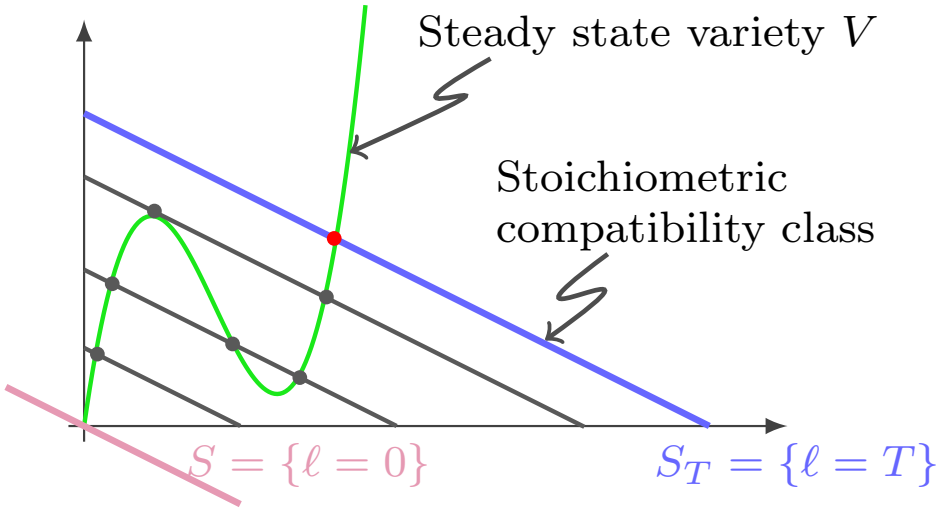


The **green curve** represents the **positive points** of the steady state variety $\mathbf{V} = \{f = 0\}$. The number of intersection points depends on the total amounts T.

If a trajectory converges, the limit is a steady state. Stable steady states attract nearby trajectories.

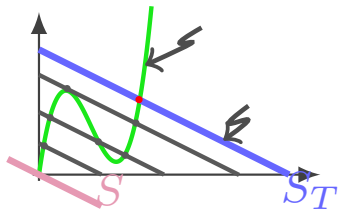
Definition

x^* is a steady state of $\frac{dx}{dt} = f(x)$ if $f(x^*) = 0$.



The **green curve** represents the **positive points** of the steady state variety $\mathbf{V} = \{f = 0\}$. The number of intersection points depends on the total amounts T.

If a trajectory converges, the limit is a steady state. Stable steady states attract nearby trajectories.



Definition

- A Chemical Reaction System (with a fix choice of κ^*) is multistationary if it is possible to find more than one positive steady state with the same T. Otherwise, we say that it is monostationary.
- A Chemical Reaction Network (with parameters κ) has the capacity of multistationarity if it is possible to find a choice of κ^* such that the associated system has more than one positive steady state with the same T.
- A region of multistationarity is an open set U in parameter space such that the CRS is multistationary for all parameters in U.

Multistationarity (multistability) is a crucial property for chemical reaction networks modeling biological processes, since its allows for different "responses" of the cell.

A non-degenerate steady state x^* is stable if $J_f(x^*)$ has $\operatorname{codim}(S)$ eigenvalues 0 and the others have negative real part.

It is expected that "half" of the steady states in the same class S_T are stable, but this requires non-trivial proofs [Hell-Rendall'15-16], [Feliu-Rendall-Wiuf'20], [Torres-Feliu'20]. Note that x^* is only given implicitly (which can be studied using the Routh-Hurwitz criterion) but we moreover have a family, a challenge.

In what follows, we will concentrate on multistationarity.

Multistationarity (multistability) is a crucial property for chemical reaction networks modeling biological processes, since its allows for different "responses" of the cell.

A non-degenerate steady state x^* is stable if $J_f(x^*)$ has $\operatorname{codim}(S)$ eigenvalues 0 and the others have negative real part.

It is expected that "half" of the steady states in the same class S_T are stable, but this requires non-trivial proofs [Hell-Rendall'15-16], [Feliu-Rendall-Wiuf'20], [Torres-Feliu'20]. Note that x^* is only given implicitly (which can be studied using the Routh-Hurwitz criterion) but we moreover have a family, a challenge.

In what follows, we will concentrate on multistationarity.

Multistationarity (multistability) is a crucial property for chemical reaction networks modeling biological processes, since its allows for different "responses" of the cell.

A non-degenerate steady state x^* is stable if $J_f(x^*)$ has $\operatorname{codim}(S)$ eigenvalues 0 and the others have negative real part.

It is expected that "half" of the steady states in the same class S_T are stable, but this requires non-trivial proofs [Hell-Rendall'15-16], [Feliu-Rendall-Wiuf'20], [Torres-Feliu'20]. Note that x^* is only given implicitly (which can be studied using the Routh-Hurwitz criterion) but we moreover have a family, a challenge.

In what follows, we will concentrate on multistationarity.

MESSI (linearly binomial) systems

- Most popular models in systems biology, in particular enzymatic networks (including the ERK pathway) have a MESSI structure [Pérez Millán-D '18]: *Modifications of type Enzyme-Substrate or Swap with Intermediates*. There is a partition of the set of species and only certain reactions can occur.
- We give combinatorial conditions on the network that ensure: 1) there are no (relevant) boundary steady states (persistence), 2) the system is conservative (and so trajectories are defined for any $t \geq 0$) and explicit equations for S are given, 3) the steady state variety V is rational, 4) V can be cut out by (explicit) binomials (and parametrized by monomials), and moreover 5) the system is *linearly* binomial [D-P. Millán-Shiu-Tang'19].
- To deal with the intermediate complexes, we can use results in [Feliu-Wiuf'13].

MESSI (linearly binomial) systems

- Most popular models in systems biology, in particular enzymatic networks (including the ERK pathway) have a MESSI structure [Pérez Millán-D '18]: *Modifications of type Enzyme-Substrate or Swap with Intermediates*. There is a partition of the set of species and only certain reactions can occur.
- We give combinatorial conditions on the network that ensure: 1) there are no (relevant) boundary steady states (persistence), 2) the system is conservative (and so trajectories are defined for any $t \geq 0$) and explicit equations for S are given, 3) the steady state variety V is rational, 4) V can be cut out by (explicit) binomials (and parametrized by monomials), and moreover 5) the system is *linearly* binomial [D-P. Millán-Shiu-Tang'19].
- To deal with the intermediate complexes, we can use results in [Feliu-Wiuf'13].

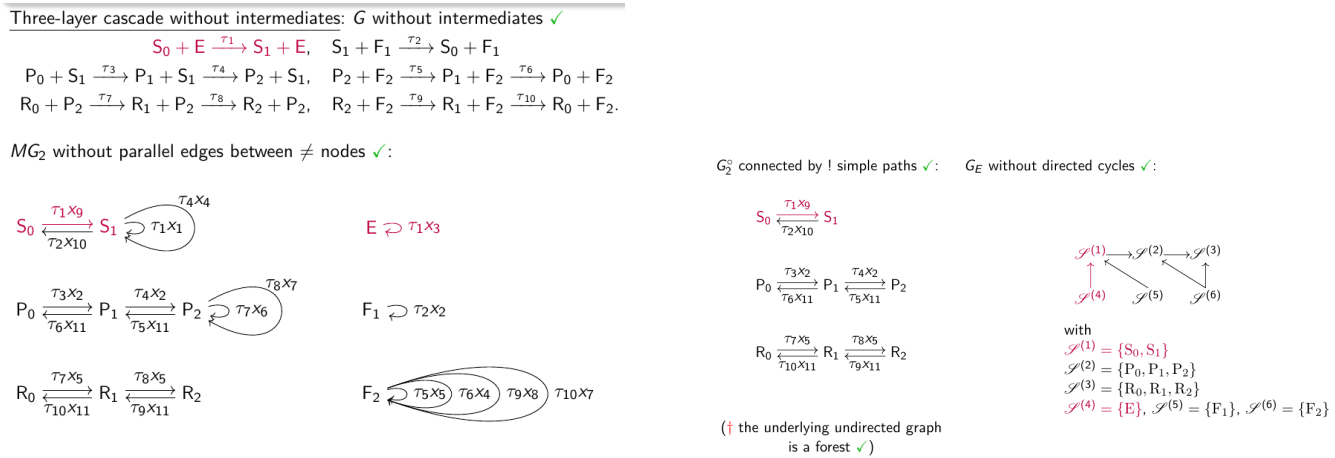
MESSI (linearly binomial) systems

- Most popular models in systems biology, in particular enzymatic networks (including the ERK pathway) have a MESSI structure [Pérez Millán-D '18]: *Modifications of type Enzyme-Substrate or Swap with Intermediates*. There is a partition of the set of species and only certain reactions can occur.
- We give combinatorial conditions on the network that ensure: 1) there are no (relevant) boundary steady states (persistence), 2) the system is conservative (and so trajectories are defined for any $t \geq 0$) and explicit equations for S are given, 3) the steady state variety V is rational, 4) V can be cut out by (explicit) binomials (and parametrized by monomials), and moreover 5) the system is *linearly* binomial [D-P. Millán-Shiu-Tang'19].
- To deal with the intermediate complexes, we can use results in [Feliu-Wiuf'13].

If we take away all the intermediates, we lose multistationarity ...

[D-Giaroli-P. Millán-Rischter '21]... to be posted

Consider a MESSI network such that some combinatorial hypotheses hold (for instance, the ERK pathway without any intermediates). Then, the associated system is monostationary for any value of (κ, T). Moreover, the system and any *canonical* extension are linearly binomial.



Questions in [Sadeghimanesh-Feliu'19]

- Where to add intermediates to the modeling to ensure the capacity for multistationarity for complete binomial networks?
- Which are the minimal subsets of intermediates with this property, termed circuits of multistationarity?

For linearly binomial networks, we ([D-G-PM-R]) implemented in Maple their criterion (based on [P. Millán-Shiu-D-Conradi'12]) by means of an equivalent formulation with a critical function (based on degree theory) which allows us to give full answers in networks like the ERK pathway or theoretically for any number of sequential phosphorylations, etc. This is beyond the capabilities of a general approach using a good CAS.

Questions in [Sadeghimanesh-Feliu'19]

- Where to add intermediates to the modeling to ensure the capacity for multistationarity for complete binomial networks?
- Which are the minimal subsets of intermediates with this property, termed circuits of multistationarity?

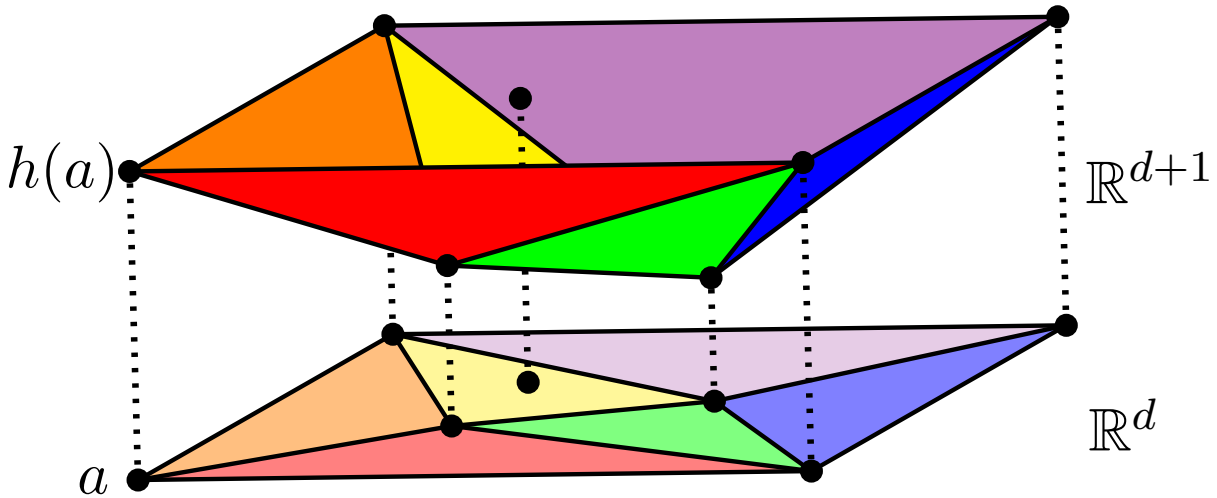
For linearly binomial networks, we ([D-G-PM-R]) implemented in Maple their criterion (based on [P. Millán-Shiu-D-Conradi'12]) by means of an equivalent formulation with a critical function (based on degree theory) which allows us to give full answers in networks like the ERK pathway or theoretically for any number of sequential phosphorylations, etc. This is beyond the capabilities of a general approach using a good CAS.

Questions in [Sadeghimanesh-Feliu'19]

- Where to add intermediates to the modeling to ensure the capacity for multistationarity for complete binomial networks?
- Which are the minimal subsets of intermediates with this property, termed circuits of multistationarity?

For linearly binomial networks, we ([D-G-PM-R]) implemented in Maple their criterion (based on [P. Millán-Shiu-D-Conradi'12]) by means of an equivalent formulation with a critical function (based on degree theory) which allows us to give full answers in networks like the ERK pathway or theoretically for any number of sequential phosphorylations, etc. This is beyond the capabilities of a general approach using a good CAS.

There is a beautiful paper by Bihan, Santos and Spaenlehauer SIAGA'18 which uses regular subdivisions of the (convex hull of the) exponents to get a lower bound on the number of positive solutions, with combinatorial arguments to get new lower bounds in terms of n and the difference between the cardinality of the support and n. This is based on classical results on degenerations and was used in [Sturmfels'94] to study real roots of complete intersections.



Example

Consider $A = \{(0,0),(2,0),(0,1),(2,1),(1,2),(1,3)\}$,

$$C = \begin{pmatrix} 1 & -2 & 1 & 1 & -1 & 0 \\ -2 & 1 & 0 & -1 & -1 & 1 \end{pmatrix}.$$

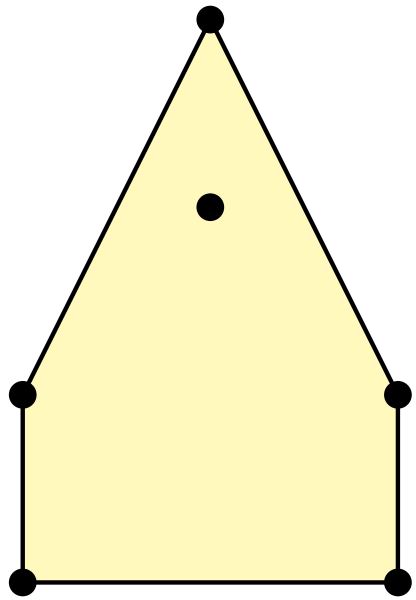
We get the polynomial system

$$\begin{array}{rcl} 1 - 2x^2 + y + x^2y - xy^2 & = & 0, \\ -2 + x^2 - x^2y - xy^2 + xy^3 & = & 0, \end{array},$$

which can be written as

$$C\begin{pmatrix} 1 & x^2 & y & x^2y & xy^2 & xy^3 \end{pmatrix}^t = 0.$$

$\mathrm{vol}_{\mathbb{Z}}(A) = 8 < 12 = 3 \cdot 4$



A definition

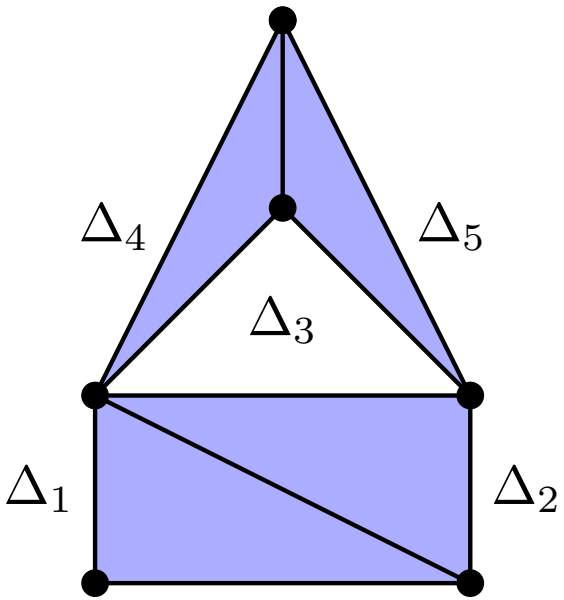
Let C be a $d \times n$ matrix with real entries. We say that a d-simplex $\Delta = \{a_{i_1}, \ldots, a_{i_{d+1}}\}$ in A is positively decorated by C if the $d \times (d+1)$ submatrix C_Δ of C with columns indicated by $\{i_1, \ldots, i_{d+1}\}$ satisfies the following:

All the values $(-1)^i \text{minor}(C_\Delta, i)$ are nonzero and have the same sign, where $\text{minor}(C_\Delta, i)$ is the determinant of the square matrix obtained by removing the i-th column.

Equivalently, all the coordinates of any non-zero vector in the kernel of the matrix C_Δ are non-zero and have the same sign.

Example

$$1 - 2\,x^2 + y + x^2y - xy^2 = 0,$$
$$-2 + x^2 - x^2y - xy^2 + xy^3 = 0,$$

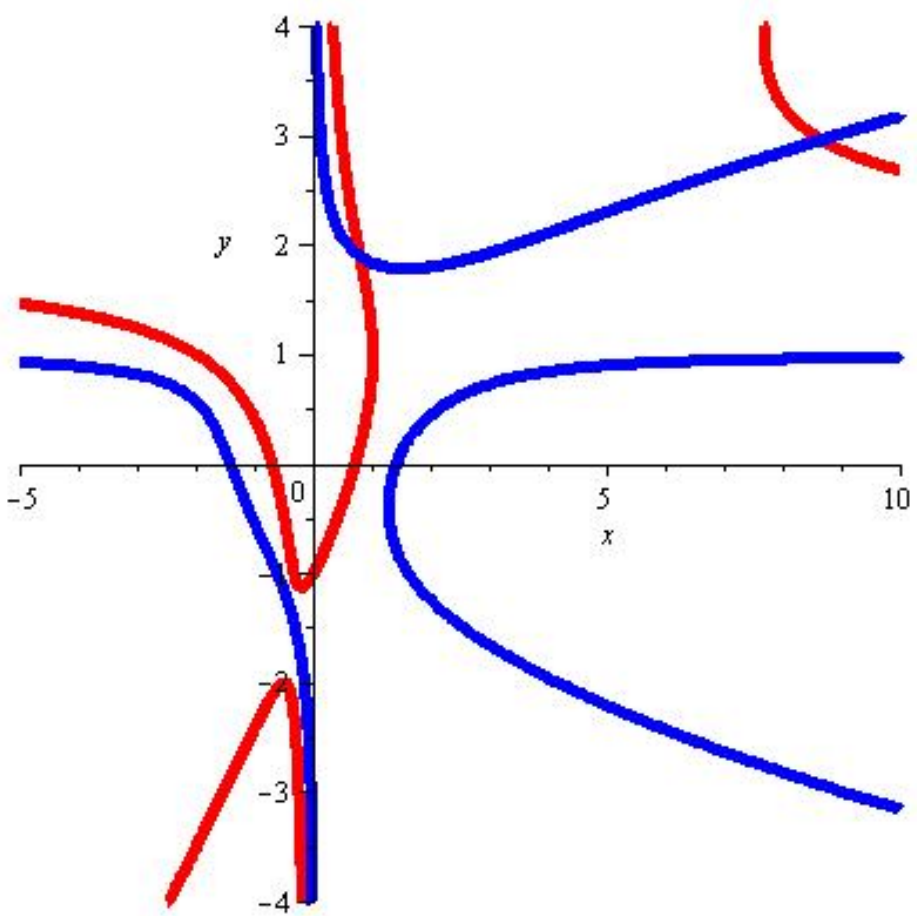


The submatrix of C given by the columns of Δ_1 (the first three columns) equals:

$$\begin{pmatrix} 1 & -2 & 1 \\ -2 & 1 & 0 \end{pmatrix}.$$

The simplex Δ_1 is positively decorated by C. Indeed, Δ_2, Δ_4 and Δ_5 are also decorated by C, but not Δ_3

$f = 0$, $g = 0$, 2 positive solutions



We can then scale/degenerate the coefficients to get a system with at least 4 positive roots . . . but we will skip this.

Degenerating with one parameter t

If we take $h \in \mathbb{R}^6$ inducing this subdivision, there exists $t_0 \in \mathbb{R}_{>0}$ such that for all $0 < t < t_0$, the number of (nondegenerate) solutions of the following deformed system is at least 4:

$$\begin{aligned} t^{h_1} - t^{h_2} 2x^2 + t^{h_3} y + t^{h_4} x^2 y - t^{h_5} xy^2 &= 0, \\ -t^{h_1} 2 + t^{h_2} x^2 - t^{h_4} x^2 y - t^{h_5} xy^2 + t^{h_6} xy^3 &= 0, \end{aligned}$$

E.g. $h_1 = 1$, $h_2 = 0$, $h_3 = 0$, $h_4 = 0$, $h_5 = 1$, $h_6 = 3$, $t = 1/12$.

- We can check e.g using a symbolic command (like `firstoct` in Singular) or numerically, that there are 4 positive roots. In general, though, the number of positively decorated simplices in a regular subdivision is smaller than the number of positive roots.
- The symbolic procedure [Pedersen-Roy-Sziprglas '91] is based on the computation of signatures of traces going back to Hermite and it doesn' t work for families (too many branchings).

Degenerating with one parameter t

If we take $h \in \mathbb{R}^6$ inducing this subdivision, there exists $t_0 \in \mathbb{R}_{>0}$ such that for all $0 < t < t_0$, the number of (nondegenerate) solutions of the following deformed system is at least 4:

$$\begin{aligned} t^{h_1} - t^{h_2} 2x^2 + t^{h_3} y + t^{h_4} x^2 y - t^{h_5} xy^2 &= 0, \\ -t^{h_1} 2 + t^{h_2} x^2 - t^{h_4} x^2 y - t^{h_5} xy^2 + t^{h_6} xy^3 &= 0, \end{aligned}$$

E.g. $h_1 = 1$, $h_2 = 0$, $h_3 = 0$, $h_4 = 0$, $h_5 = 1$, $h_6 = 3$, $t = 1/12$.

- We can check e.g using a symbolic command (like `firstoct` in Singular) or numerically, that there are 4 positive roots. In general, though, the number of positively decorated simplices in a regular subdivision is smaller than the number of positive roots.
- The symbolic procedure [Pedersen-Roy-Sziprglas '91] is based on the computation of signatures of traces going back to Hermite and it doesn' t work for families (too many branchings).

Degenerating with one parameter t

If we take $h \in \mathbb{R}^6$ inducing this subdivision, there exists $t_0 \in \mathbb{R}_{>0}$ such that for all $0 < t < t_0$, the number of (nondegenerate) solutions of the following deformed system is at least 4:

$$\begin{aligned} t^{h_1} - t^{h_2} 2x^2 + t^{h_3} y + t^{h_4} x^2 y - t^{h_5} xy^2 &= 0, \\ -t^{h_1} 2 + t^{h_2} x^2 - t^{h_4} x^2 y - t^{h_5} xy^2 + t^{h_6} xy^3 &= 0, \end{aligned}$$

E.g. $h_1 = 1$, $h_2 = 0$, $h_3 = 0$, $h_4 = 0$, $h_5 = 1$, $h_6 = 3$, $t = 1/12$.

- We can check e.g using a symbolic command (like `firstoct` in Singular) or numerically, that there are 4 positive roots. In general, though, the number of positively decorated simplices in a regular subdivision is smaller than the number of positive roots.
- The symbolic procedure [Pedersen-Roy-Sziprglas '91] is based on the computation of signatures of traces going back to Hermite and it doesn' t work for families (too many branchings).

Degenerating with one parameter t

If we take $h \in \mathbb{R}^6$ inducing this subdivision, there exists $t_0 \in \mathbb{R}_{>0}$ such that for all $0 < t < t_0$, the number of (nondegenerate) solutions of the following deformed system is at least 4:

$$\begin{aligned} t^{h_1} - t^{h_2} 2x^2 + t^{h_3} y + t^{h_4} x^2 y - t^{h_5} xy^2 &= 0, \\ -t^{h_1} 2 + t^{h_2} x^2 - t^{h_4} x^2 y - t^{h_5} xy^2 + t^{h_6} xy^3 &= 0, \end{aligned}$$

E.g. $h_1 = 1$, $h_2 = 0$, $h_3 = 0$, $h_4 = 0$, $h_5 = 1$, $h_6 = 3$, $t = 1/12$.

- We can check e.g using a symbolic command (like `firstoct` in Singular) or numerically, that there are 4 positive roots. In general, though, the number of positively decorated simplices in a regular subdivision is smaller than the number of positive roots.
- The symbolic procedure [Pedersen-Roy-Sziprglas '91] is based on the computation of signatures of traces going back to Hermite and it doesn' t work for families (too many branchings).

Obtaining a region of multistationarity

[Bihan, D., Giaroli, J. Algebra' 20]

Let $A = \{a_1, \dots, a_N\}$ in $\mathbb{R}^n$ and $C = (c_{i,j}) \in \mathbb{R}^{n \times N}$. Assume there are p n-simplices $\Delta_1, \dots, \Delta_p$ contained in A, which are part of a regular subdivision of A and are positively decorated by C.

Let $\mathcal{C}_{\Delta_1,\dots,\Delta_p}$ be the cone of all height vectors $h \in \mathbb{R}^N$ that induce a regular subdivision of A containing $\Delta_1, \dots, \Delta_p$:

$$\mathcal{C}_{\Delta_1,\dots,\Delta_p} = \{h \in \mathbb{R}^N : \langle m_r, h \rangle > 0,\ r = 1, \dots, \ell\}. \tag{1}$$

Then, $\forall \varepsilon \in (0,1)^\ell$ there exists $t_0(\varepsilon) > 0$ s.t $\forall \gamma$ in the open set $\mathbf{U}$

$$\mathbf{U} = \cup_{\varepsilon \in (0,1)^\ell} \{\gamma \in \mathbb{R}^N_{>0} ;\ \gamma^{m_r} < t_0(\varepsilon)^{\varepsilon_r},\ r = 1 \dots, \ell\},$$

the system

$$\sum_{j=1}^{n} c_{ij}\, \gamma_j\, x^{a_j} = 0, \quad i = 1, \dots, n, \tag{2}$$

has at least p nondegenerate positive solutions.

Obtaining a region of multistationarity

[Bihan, D., Giaroli, J. Algebra' 20]

Let $A = \{a_1, \dots, a_N\}$ in $\mathbb{R}^n$ and $C = (c_{i,j}) \in \mathbb{R}^{n \times N}$. Assume there are p n-simplices $\Delta_1, \dots, \Delta_p$ contained in A, which are part of a regular subdivision of A and are positively decorated by C.
Let $\mathcal{C}_{\Delta_1,\dots,\Delta_p}$ be the cone of all height vectors $h \in \mathbb{R}^N$ that induce a regular subdivision of A containing $\Delta_1, \dots, \Delta_p$:

$$\mathcal{C}_{\Delta_1,\dots,\Delta_p} = \{h \in \mathbb{R}^N : \langle m_r, h \rangle > 0,\ r = 1, \dots, \ell\}. \tag{1}$$

Then, $\forall \varepsilon \in (0,1)^\ell$ there exists $t_0(\varepsilon) > 0$ s.t $\forall \gamma$ in the open set $\mathbf{U}$

$$\mathbf{U} = \cup_{\varepsilon \in (0,1)^\ell} \{\gamma \in \mathbb{R}^N_{>0} ;\ \gamma^{m_r} < t_0(\varepsilon)^{\varepsilon_r},\ r = 1 \dots, \ell\},$$

the system

$$\sum_{j=1}^{n} c_{ij}\, \gamma_j\, x^{a_j} = 0, \quad i = 1, \dots, n, \tag{2}$$

has at least p nondegenerate positive solutions.

Obtaining a region of multistationarity

[Bihan, D., Giaroli, J. Algebra' 20]

Let $A = \{a_1, \dots, a_N\}$ in $\mathbb{R}^n$ and $C = (c_{i,j}) \in \mathbb{R}^{n \times N}$. Assume there are p n-simplices $\Delta_1, \dots, \Delta_p$ contained in A, which are part of a regular subdivision of A and are positively decorated by C.
Let $\mathcal{C}_{\Delta_1,\dots,\Delta_p}$ be the cone of all height vectors $h \in \mathbb{R}^N$ that induce a regular subdivision of A containing $\Delta_1, \dots, \Delta_p$:

$$\mathcal{C}_{\Delta_1,\dots,\Delta_p} = \{h \in \mathbb{R}^N : \langle m_r, h \rangle > 0,\ r = 1, \dots, \ell\}. \tag{1}$$

Then, $\forall \varepsilon \in (0,1)^\ell$ there exists $t_0(\varepsilon) > 0$ s.t $\forall \gamma$ in the open set $\mathbf{U}$

$$\mathbf{U} = \cup_{\varepsilon \in (0,1)^\ell} \{\gamma \in \mathbb{R}^N_{>0} \,;\, \gamma^{m_r} < t_0(\varepsilon)^{\varepsilon_r},\ r = 1 \dots, \ell\},$$

the system

$$\sum_{j=1}^{n} c_{ij}\, \gamma_j\, x^{a_j} = 0, \quad i = 1, \dots, n, \tag{2}$$

has at least p nondegenerate positive solutions.

Difficulties we need to overcome

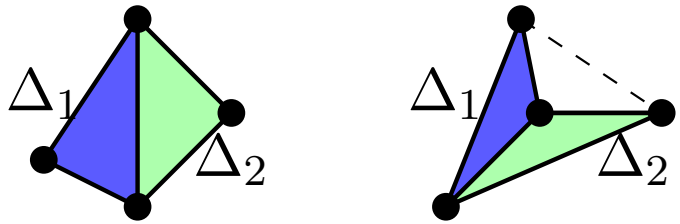
- Even if deciding if simplices are part of a same regular subdivision is algorithmic, how to do this when the dimension or the number of monomials is big (or when they are not upper bounded)?

One way out: If two simplices share a facet, then this is always the case! But this restricts our lower bound to 2... in fact, to 3! We were able to find more for sequential phosphorylations with n-sites[G-R-PM-D. '19]

- We get polynomials with non-generic coefficients, which are rational functions of the original rate constants κ and need to assert that we can rescale κ. We heavily use the results about the structure of (s-toric) MESSI systems.

Difficulties we need to overcome

- Even if deciding if simplices are part of a same regular subdivision is algorithmic, how to do this when the dimension or the number of monomials is big (or when they are not upper bounded)?

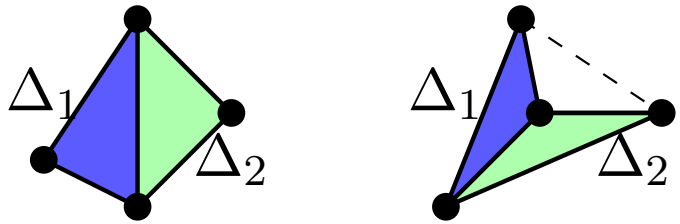


One way out: If two simplices share a facet, then this is always the case! But this restricts our lower bound to 2... in fact, to 3! We were able to find more for sequential phosphorylations with n-sites[G-R-PM-D. '19]

- We get polynomials with non-generic coefficients, which are rational functions of the original rate constants κ and need to assert that we can rescale κ. We heavily use the results about the structure of (s-toric) MESSI systems.

Difficulties we need to overcome

- Even if deciding if simplices are part of a same regular subdivision is algorithmic, how to do this when the dimension or the number of monomials is big (or when they are not upper bounded)?



  One way out: If two simplices share a facet, then this is always the case! But this restricts our lower bound to 2... in fact, to 3! We were able to find more for sequential phosphorylations with n-sites[G-R-PM-D. '19]
- We get polynomials with non-generic coefficients, which are rational functions of the original rate constants κ and need to assert that we can rescale κ. We heavily use the results about the structure of (s-toric) MESSI systems.

Other current computational approaches

- Methods using degree theory

  [Conradi-Feliu-Mincheva-Wiuf'17]

- Symbolic software

  GB, RAGlib, CAD software, etc.

- Tropical tools to separate time scales and SMT solvers

  [Kruff-Lüders-Radulescu-Sturm-Walcher'21]

- SONC and other tools

  [Feliu-Kaihnsa-de Wolff-Yürük'20-21] (n-site)

- Numerical algebraic geometry

  [Gross-Harrington-Rosen-Sturmfels'16],
  [Nam-Gyori-Amethyst-Bates-Gunawardena'20]

Future computational approaches?

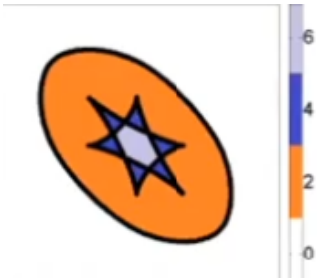
- Massive parallel computations in AG

[Böhm-Decker-Frühbis-Krüger-Pfreundt-Rahn-Ristau '21]

- Machine learning to improve GB computations

[Peifer-Stillman-Halpern Leistner'21]

- Machine learning to approximate the real discriminant

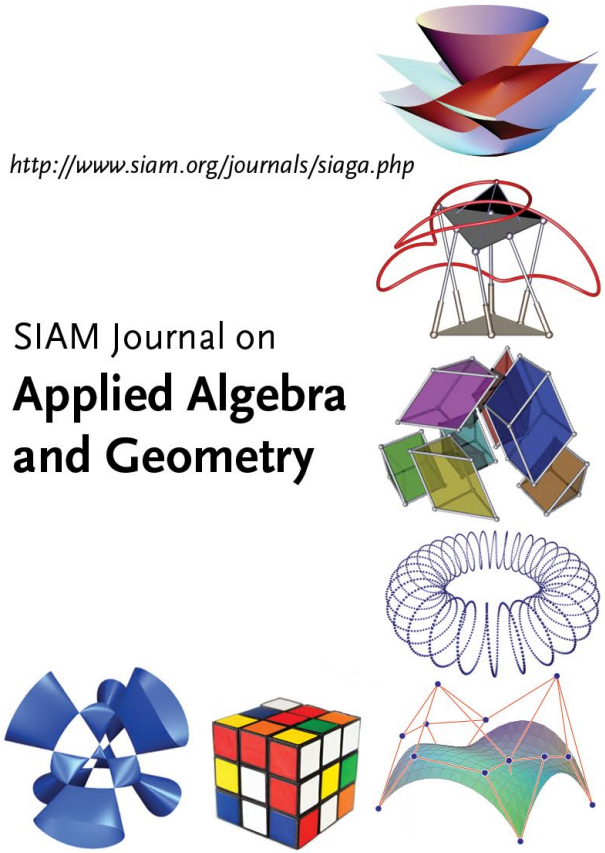


[Bernal-Hauenstein-Mehta-Regan-Tang'20],

- Machine learning to unveil structure, to discover math theorems

[He'21]

♡ Many thanks for your attention! ♡



UNION MATEMATICA ARGENTINA
FUNDADA EN 1936

Revista de la Unión Matemática Argentina is a **diamond open access** journal that publishes original articles in all areas of pure and applied mathematics.

Editor-in-chief: Alicia Dickenstein

Impact factor for 2017: 0.517 (Journal Citation Reports)

Publication frequency: two issues per year

There are no charges for authors and readers. On the journal's homepage you may find the complete Editorial Board, submission information, and the full text of all articles since the journal's foundation in 1936.

http://inmabb.criba.edu.ar/revuma/