REALIZABILITY IN CHEMICAL REACTION NETWORK THEORY

OSKAR HENRIKSSON

1. Preliminaries

A reaction network with species $\{X_1, \dots, X_n\}$ is a list of formal expressions of the form

$$\begin{aligned} a_{11}X_1 + a_{21}X_2 + \cdots + a_{n1}X_n &\longrightarrow b_{11}X_1 + b_{21}X_2 + \cdots + b_{n1}X_n \\ &\vdots \\ a_{1r}X_1 + a_{2r}X_2 + \cdots + a_{nr}X_n &\longrightarrow b_{1r}X_1 + b_{2r}X_2 + \cdots + b_{nr}X_n\,. \end{aligned}$$

Since the species do not have mathematical significance, all the relevant data of the network is contained in the so-called ***reactant matrix*** $A = (a_{ij}) \in \mathbb{Z}_{\geq 0}^{n\times r}$ and ***product matrix*** $B = (b_{ij}) \in \mathbb{Z}_{\geq 0}^{r\times n}$. Here, n is the number of species of the network, and r the number of reactions.

Under the assumption of ***mass-action kinetics***, a network given by matrices $A, B \in \mathbb{Z}_{\geq 0}^{n\times r}$ gives rise to a tuple of parametric ***steady state polynomials***

$$F_{A,B}(\boldsymbol{\kappa}, \boldsymbol{x}) := (B - A)(\boldsymbol{\kappa} \circ \boldsymbol{x}^A) \in \mathbb{Q}[\kappa_1, \dots, \kappa_r, x_1, \dots, x_n]^n\,, \tag{1.1}$$

where $\boldsymbol{x} = (x_1, \dots, x_n)$ are considered as variables (representing the concentration of the species), $\boldsymbol{\kappa} = (\kappa_1, \dots, \kappa_r)$ are considered as parameters (often referred to as rate constants), and $\circ$ denotes componentwise multiplication. For each choice of rate constants $\boldsymbol{\kappa}^* \in \mathbb{R}_{>0}^r$, we get a tuple of ***specialized steady state polynomials***

$$F_{A,B,\boldsymbol{\kappa}^*}(\boldsymbol{x}) := F_{A,B}(\boldsymbol{\kappa}^*, \boldsymbol{x}) \in \mathbb{R}[x_1, \dots, x_n]^n\,.$$

Example 1.1. The reaction network

$$\begin{aligned} 2X_1 + X_4 &\longrightarrow 3X_2 + 2X_4 \\ 3X_2 + 2X_4 &\longrightarrow 4X_3 \\ 4X_3 &\longrightarrow 2X_1 + X_4 \end{aligned} \tag{1.2}$$

can be encoded by the matrices

$$A = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 4 \\ 1 & 2 & 0 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 0 & 0 & 2 \\ 3 & 0 & 0 \\ 0 & 4 & 0 \\ 2 & 0 & 1 \end{bmatrix},$$

and gives rise to the following steady state polynomials:

$$F_{A,B}(\boldsymbol{\kappa}, \boldsymbol{x}) = \begin{bmatrix} -2\kappa_1 x_1^2 x_4 + 2\kappa_3 x_3^4 \\ 3\kappa_1 x_1^2 x_4 - 3\kappa_2 x_2^3 x_4^2 \\ 4\kappa_2 x_2^3 x_4^2 - 4\kappa_3 x_3^4 \\ \kappa_1 x_1^2 x_4 - 2\kappa_2 x_2^3 x_4^2 + \kappa_3 x_3^4 \end{bmatrix}. \tag{1.3}$$

2. Realizability

A common theme in reaction network theory is to ask questions of the following type.

Question 2.1. *Given a polynomial system $G \in \mathbb{R}[x_1, \dots, x_n]^n$, can we find a reaction network, encoded by matrices $A, B \in \mathbb{Z}_{\geq 0}^{n\times r}$ for some $r \in \mathbb{Z}_{>0}$, and a choice of rate constants $\boldsymbol{\kappa}^* \in \mathbb{R}_{>0}^r$ that together* **realize** *G, in the sense that $G = F_{A,B,\boldsymbol{\kappa}^*}$?*

There are many possible variations of Question 2.1 that are of interest. For instance:

- Realizability by particular classes of networks (e.g., weakly reversible ones).
- Realizability of coarser objects defined by polynomials, e.g. ideals, varieties or semialgebraic sets.
- Realizability of whole families of polynomials $(G_\alpha)_{\alpha\in\mathcal{A}}$ as subfamilies of families of the form $(f_{A,B,\boldsymbol{\kappa}^*})_{\boldsymbol{\kappa}^*\in\mathbb{R}^r_{>0}}$.

In addition, there are also associated questions of uniqueness and optimality.

One of the most well-known realizability results is due to Hárs and Tóth, and is often referred to as the "Hungarian Lemma".

Theorem 2.2 ([HT79, Thm. 3.2])**.** *Consider a tuple of polynomials $G = C\boldsymbol{x}^M \in \mathbb{R}[x_1,\dots,x_n]^n$, with m distinct monomials encoded as the columns of a matrix $M \in \mathbb{Z}^{n\times m}_{\geq 0}$, and coefficient matrix $C \in \mathbb{R}^{n\times m}$. Then there exists $r \in \mathbb{Z}_{>0}$, $A, B \in \mathbb{Z}^{n\times r}_{\geq 0}$ and $\boldsymbol{\kappa}^* \in \mathbb{R}^r_{>0}$ such that $G = F_{A,B,\boldsymbol{\kappa}^*}$ if and only if $M_{ij} > 0$ whenever $C_{ij} < 0$.*

Proof. The "only if" direction is immediate from (1.1). The proof of the "if" direction given in [HT79] is constructive, via Algorithm 2.3. □

Algorithm 2.3.

Input: A square polynomial system in the form of an exponent matrix $M \in \mathbb{Z}^{n\times m}_{\geq 0}$, and a coefficient matrix $C \in \mathbb{R}^{n\times m}$

Output: A realization in the form of matrices $A, B \in \mathbb{Z}^{n\times r}_{\geq 0}$ for some $r \in \mathbb{Z}_{>0}$, and a choice of rate constants $\boldsymbol{\kappa}^* \in \mathbb{R}^r_{>0}$ such that $C\boldsymbol{x}^M = F_{A,B,\boldsymbol{\kappa}^*}$

Initiate empty matrices A and B, and an empty vector $\boldsymbol{\kappa}^*$
for $i \in [n]$ **do**
  for $j \in [m]$ **do**
    if $C_{ij} \neq 0$ **then**
      # Create a reaction that contributes a term $C_{ij}\boldsymbol{x}^{M_{*j}}$ to the ith polynomial
      # without affecting the other polynomials
      Append the column M_{*j} to A
      Append the column $M_{*j} + \operatorname{sign}(C_{ij})\boldsymbol{e}_i$ to B ($\boldsymbol{e}_i$ denotes the ith standard basis vector)
      Append the entry $|C_{ij}|$ to $\boldsymbol{\kappa}^*$
return $A, B, \boldsymbol{\kappa}^*$

The output of Algorithm 2.3 is called the ***canonical realization*** of the system G. The number of reactions in it will simply be the sum of the number of terms in the respective polynomials. It is worth noting that this is typically far from the only possible realization, and that other realizations might be more economical to work with, or biologically more meaningful.

Example 2.4. The following parametric system (which describes the steady states of an epidemiological SIR model, with X_1, X_2 and X_3 representing susceptible, infected and recovered individuals, respetively) satisfies the sign condition in Theorem 2.2 for any parameters $\beta, \gamma \in \mathbb{R}_{>0}$:

$$G = [-\beta x_1 x_2, \beta x_1 x_2 - \gamma x_2, \gamma x_2]\,,$$

and the canonical realization is given by the following network with $\boldsymbol{\kappa}^* = (\beta, \beta, \gamma, \gamma)$:

$$\begin{aligned} X_1 + X_2 &\longrightarrow X_2 \\ X_1 + X_2 &\longrightarrow 2X_2 \\ X_2 &\longrightarrow 0 \\ X_2 &\longrightarrow X_2 + X_3\,. \end{aligned}$$

Another (more biologically meaningful) realization is the following network with $\boldsymbol{\kappa}^* = (\beta, \gamma)$:

$$\begin{aligned} X_1 + X_2 &\longrightarrow 2X_2 \\ X_2 &\longrightarrow X_3\,. \end{aligned}$$

Example 2.5 ([HT79, Ex. 3.1])**.** If we apply the algorithm to the system (1.3) with rate constants $\tilde{\boldsymbol{\kappa}} \in \mathbb{R}^3_{>0}$, the canonical realization turns out to be a different network than (1.2), namely:

$$\begin{aligned} 2X_1 + X_4 &\longrightarrow X_1 + X_4 \\ 4X_3 &\longrightarrow X_1 + 4X_3 \\ 2X_1 + X_4 &\longrightarrow 2X_1 + X_2 + X_4 \\ 3X_2 + 2X_4 &\longrightarrow 2X_2 + 2X_4 \\ 3X_2 + 2X_4 &\longrightarrow 3X_2 + X_3 + 2X_4 \\ 4X_3 &\longrightarrow 3X_3 \\ 2X_1 + X_4 &\longrightarrow 2X_1 + 2X_4 \\ 3X_2 + 2X_4 &\longrightarrow 3X_2 + X_4 \\ 4X_3 &\longrightarrow 4X_3 + X_4 \end{aligned}$$

with the following choice of rate constants:

$$\boldsymbol{\kappa}^* = (2\tilde{\kappa}_1, 2\tilde{\kappa}_3, 3\tilde{\kappa}_1, 3\tilde{\kappa}_2, 4\tilde{\kappa}_2, 4\tilde{\kappa}_3, \tilde{\kappa}_1, 2\tilde{\kappa}_2, \tilde{\kappa}_3)\ .$$

Example 2.6. A classical example of an unrealizable parametric system (which one might recognize as the right-hand side of a Lorentz attractor system) is

$$G = [\sigma x_2 - \sigma x_1, \rho x_1 - x_1 x_3 - x_2, x_1 x_2 - \beta x_3]\,, \tag{2.1}$$

where the second monomial in the second polynomial violates the sign condition of Theorem 2.2 for any $\sigma, \rho, \beta \in \mathbb{R}_{>0}$.

We now turn to the question of realizing "coarser" objects that depend on a polynomial system, such as ideals and solution sets. The following is an almost immediate consequence of Theorem 2.2 (see for instance [Dic16, Sect. 2] for a short discussion in this direction).

Corollary 2.7. *Let $h_1, \dots, h_s \in \mathbb{R}[x_1, \dots, x_n]$ for $s \leq n$ be arbitrary polynomials. Then there exists $r \in \mathbb{Z}_{>0}$, $A, B \in \mathbb{Z}^{n\times r}$ and $\boldsymbol{\kappa}^* \in \mathbb{R}^r_{>0}$ such that the following equalities hold:*

(i) $\langle h_1, \dots, h_s\rangle = \langle F_{A,B,\boldsymbol{\kappa}^*}\rangle$ *as ideals in the Laurent polynomial ring* $\mathbb{R}[x_1^{\pm}, \dots, x_n^{\pm}]$
(ii) $\mathbb{V}(h_1, \dots, h_s) \cap (\mathbb{R}^*)^n = \mathbb{V}(F_{A,B,\boldsymbol{\kappa}^*}) \cap (\mathbb{R}^*)^n$
(iii) $\mathbb{V}(h_1, \dots, h_s) \cap \mathbb{R}^n_{>0} = \mathbb{V}(F_{A,B,\boldsymbol{\kappa}^*}) \cap \mathbb{R}^n_{>0}$.

Proof. Form a square polynomial system $G = (g_1, \dots, g_n) \in \mathbb{R}[x_1, \dots, x_n]^n$ by setting

$$g_1 = x_1 h_1\,, \quad \dots\,, \quad g_s = x_s h_s\,, \quad g_{s+1} = 0\,, \quad \dots\,, \quad g_n = 0\,,$$

and apply Theorem 2.2 to find a realization $F_{A,B,\boldsymbol{\kappa}^*} = G$. Part (i) then follows by noting that $\langle h_1, \dots, h_s\rangle = \langle G\rangle$ in $\mathbb{R}[x_1^{\pm}, \dots, x_n^{\pm}]$, whereas parts (ii) and (iii) follow directly from (i). □

Example 2.8. The positive roots of (2.1) coincide with the positive real roots of $[\sigma x_2 - \sigma x_1, \rho x_1 x_2 - x_1 x_2 x_3 - x_2^2, x_1 x_2 - \beta x_3]$, and can thus be realized as $\mathbb{V}(F_{A,B,\boldsymbol{\kappa}^*}) \cap \mathbb{R}^3_{>0}$ for the following network, with $\boldsymbol{\kappa}^* = (\sigma, \sigma, \rho, 1, 1, 1, \beta)$:

$$\begin{aligned} X_2 &\longrightarrow X_2 + X_1 \\ X_1 &\longrightarrow 0 \\ X_1 + X_2 &\longrightarrow X_1 + 2X_2 \\ X_1 + X_2 + X_3 &\longrightarrow X_1 + X_3 \\ 2X_2 &\longrightarrow X_2 \\ X_1 + X_2 &\longrightarrow X_1 + X_2 + X_3 \\ X_3 &\longrightarrow 0\,. \end{aligned}$$

We end with the observation that part (ii) and (iii) of Corollary 2.7 hold also for $s > n$, since we always have

$$\mathbb{V}_{\mathbb{R}}(h_1, \ldots, h_s) = \mathbb{V}_{\mathbb{R}}(h_1, \ldots, h_n^2 + h_{n+1}^2 + \cdots + h_s^2)\,.$$

We collect our conclusions in a final result.

Theorem 2.9 (The extended "Hungarian lemma")**.** *Let $h_1, \ldots, h_s \in \mathbb{R}[x_1, \ldots, x_n]$ be arbitrary polynomials for any integer $s > 0$. Then there exists $r \in \mathbb{Z}_{>0}$, $A, B \in \mathbb{Z}^{n\times r}$ and $\boldsymbol{\kappa}^* \in \mathbb{R}^r_{>0}$ such that the following equalities hold:*

(i) $\mathbb{V}(h_1, \ldots, h_s) \cap (\mathbb{R}^*)^n = \mathbb{V}(F_{A,B,\boldsymbol{\kappa}^*}) \cap (\mathbb{R}^*)^n$
(ii) $\mathbb{V}(h_1, \ldots, h_s) \cap \mathbb{R}^n_{>0} = \mathbb{V}(F_{A,B,\boldsymbol{\kappa}^*}) \cap \mathbb{R}^n_{>0}$.

References

[Dic16] A. Dickenstein. Biochemical reaction networks: an invitation for algebraic geometers. In *Mathematical Congress of the Americas 2013*, volume 656 of *Contemporary Mathematics*, page 65–83. American Mathematical Society, 2016. Cited on page 3.

[HT79] V. Hárs and J. Tóth. On the inverse problem of reaction kinetics. In *Colloquia Mathematica Societatis János Bolyai 30., Qualitative Theory of Differential Equations, Szeged (Hungary)*, pages 363–379, 1979. Cited on pages 2 and 3.