What are... some graphs of chemical reaction networks?

M. Domijan
Mathematics Institute, University of Warwick, Coventry CV4 7AL, United Kingdom
February 4, 2008

Abstract

In order to make sense of biological processes taking part on a cellular level, it is crucial to understand the underlying chemical reactions. In recent years, graph theory has emerged as one of the most approachable and fruitful methods for analysing chemical reaction networks (CRNs). Graph theory can deal with models for which other techniques fail, for example, models where there is incomplete information about the parameters or that are of high dimension. Models with such issues are common in CRN theory. We will start with elementary definitions and move on to review some of the most recent graph theory that has been applied to studies of CRNs.

1 Introduction

Chemical reaction networks (CRNs) are a rich area of application for mathematical theory. One of the main reasons is that various CRNs can exhibit strikingly different dynamic behaviour, from multiple steady states to stable oscillations [2, 36, 46, 32, 37, 4, 22, 21, 49, 44, 45, 29]. Their diverse behaviour arises from the experimental set-up, whereby the reactions and their environment can be strictly controlled. For example, do the reactions take place in a closed environment where none of the species have inflow or outflow, or in an open environment where at least one of the species can enter or leave the system?

A great deal of the qualitative studies into the chemical reaction networks have been amassed since the early work of the 70s by Bruce Clarke, Martin Feinberg, Horn and Jackson, Ivanova and others, see [6, 7, 8, 18, 19, 30, 31, 33, 34, 59]. These authors have worked to develop analytic tools for studying dynamic behaviour of CRNs that can overcome several of the issues that are frequent in mathematical modeling.

One of the important issues is the lack of information about parameters in chemical reaction models. For models reviewed in this paper it is assumed that reactions obey the law of mass action. The models describe the dynamics of species concentrations. Species concentrations depend on the rates of reactions in which the species participate. If reactions obey law of mass-action, reaction rates depend
on a product of the concentrations of all the reacting species and a coefficient, called kinetic coefficient, that combines other effects on the reaction such as temperature, light and ionic bonds of species. Each reaction has at least one kinetic constant associated to it. When a model for the species dynamics is derived, the kinetic coefficients become the parameters of the system of ODEs. Normally it is difficult to measure these coefficients in experiments, hence much chemical literature is devoted to finding ways to analyse system dynamics by avoiding the use of the kinetic coefficients.

Another important issue is the size of the network. Size of the network is the number of species taking part in the chemical reactions. While many methods exist for analysing small networks, they may be computationally difficult to implement on networks with several hundred species.

Graph theory is lauded as an approach that can successfully deal with both of the aforementioned issues \[3, 50\]. Chemical reaction networks can be associated to various types of graphs. For some examples reader should refer to \[53, 18, 31, 5, 10, 41, 48\]. In this paper, a selection of three graphs will be reviewed. All three graphs have recently appeared in literature: the interaction graph, the species-reaction graph and a relative of the species reaction graph, called the directed bipartite graph. Their structures have been known for a long time, but there has been much recent development relating them to dynamic behaviour of CRNs. Aside from introducing the graphs, where it is appropriate comparisons between them will be drawn. There appears to be an underlying trade-off between the amount of system information a graph should contain and the knowledge of the system dynamics that the graph can give. Ideally one would like to design a simple graph that requires the least amount of information, but from which one can still draw meaningful conclusions. It is clearly important that the understanding of the connections between the three graphs will help with further development of graph theoretic methods for analysis of CRNs.

2 Background

A chemical reaction system with \(m\) reactions and \(n\) reacting species is represented by time-continuous dynamical systems that are derived from reaction schemes. A typical set of reactions is written in a chemical reaction scheme,

\[
\sum_{i=1}^{n} \alpha_{ji} A_i \rightarrow \sum_{i=1}^{n} \beta_{ji} A_i, \quad j = 1, ..., m
\]  

(1)

where \(A_i, 1 \leq i \leq n\), are the reacting species participating in \(j\)-th reaction. The coefficients \(\alpha_{ji}\) and \(\beta_{ji}\) represent the number of \(A_i\) molecules participating in \(j\)-th reaction at reactant and product stages, respectively.

The objects at the head or tail of the arrow in reaction scheme are called complexes. Let us define \(\mathbb{R}_+\) to be the set of positive real numbers and \(\mathbb{R}_n\) to be the set of nonnegative real numbers. Taking the formal vector space with the species set \(\{A_1, \ldots, A_n\}\) as a basis, each complex can be associated naturally with a formal sum. A reaction complex is the set of species at the tail of the arrow, labeled by a formal sum \(\sum_{i=1}^{n} \alpha_{ji} A_i\) and a product complex is set of species at the head of the arrow, labeled \(\sum_{i=1}^{n} \beta_{ji} A_i\). Likewise, the complexes can be written in
vector form, for example, \( y_j = (\alpha_{j1}, \ldots, \alpha_{jn})^T \). Labeling the index set of species \( \mathcal{S} \), it becomes clear that \( y_j \in \mathbb{R}^S_+ \).

The chemical reaction network (CRN) can be derived from the chemical reaction scheme:

**Definition 2.1.** A chemical reaction network \( \mathcal{N} = (\mathcal{S}, \mathcal{C}, \mathcal{R}) \) consists of three sets:

1. a set of network species \( \mathcal{S} \),
2. a set of the network complexes \( \mathcal{C} \subset \mathbb{R}^S_+ \),
3. a set of reactions \( \mathcal{R} \subset \mathcal{C} \times \mathcal{C} \) with
   - (a) \((y, y) \notin \mathcal{R}\) for any \( y \in \mathcal{C} \) and
   - (b) for every \( y \in \mathcal{C} \) there exists a \( y' \in \mathcal{C} \) such that \((y, y') \in \mathcal{R}\) or \((y', y) \in \mathcal{R}\).

When \((y, y') \in \mathcal{R}\), complex \( y \) reacts to complex \( y' \). Such a reaction is denoted by \( y \rightarrow y' \). If \( \{(y, y'), (y', y)\} \subset \mathcal{R} \) the set \( \{y \rightarrow y', y' \rightarrow y\} \) is denoted by \( y \mathrel{\leftrightarrow} y' \). Reaction \( y \mathrel{\leftrightarrow} y' \) is a reversible reaction. On the other hand, if \((y, y') \in \mathcal{R}\) and \((y', y) \notin \mathcal{R}\), reaction \( y \rightarrow y' \) is irreversible.

There are two special types of reactions that require mention. A reaction is called an inflow reaction if the reaction complex vector is a zero vector, namely \( 0 \rightarrow y' \). It is called an outflow reaction if the product complex vector is a zero vector, \( y \rightarrow 0 \). Inflow and outflow reactions model the interaction of the reaction system with the outside environment. Systems can be classed according to their interaction with external species. A reaction system is called open, if it has at least one inflow or an outflow reaction. On the other hand, the system is closed if it has no inflow nor outflow reactions.

This paper assumes that modeling CRNs takes a deterministic approach, taking as an assumption that all species come in large numbers and hence their concentrations are continuous. Mass-action kinetics is one of the possible types of kinetics under which chemical species interact. They are associated with polynomial ODEs, that shall now be described.

Consider a reaction
\[
\alpha_{j1}A_1 + \alpha_{j2}A_2 + \alpha_{j3}A_3 \rightarrow \beta_{j1}A_1 + \beta_{j4}A_4.
\]
It has the reaction rate
\[
v_j = k \times [A_1]^\alpha_{j1} \times [A_2]^\alpha_{j2} \times [A_3]^\alpha_{j3}
\]
where \( k \) is kinetic coefficient and \([A_j]\) is the concentration of species \( A_j \) (i.e., number of molecules of \( A_j \) in a given volume). Dynamics of \([A_2]\) depend on rate \( v_j \) in the form,

\[
\dot{[A_2]} = (\beta_{j2} - \alpha_{j2})v_j = (-\alpha_{j2})v_j
\]

Assuming the system has \( i \) irreversible and \( r \) reversible reactions, under mass-action kinetics, \( p = i + 2 \times r \) kinetic coefficients will be associated with the set of reactions. A mass-action system is a CRN \((\mathcal{S}, \mathcal{C}, \mathcal{R})\) with the kinetic coefficients \( k \in \mathbb{R}^p_+ \).
Define $k_{y_j \rightarrow y'_j}$ to be the kinetic coefficient of reaction $y_j \rightarrow y'_j$ and relabel $[A_i]$ by $x_i$. For $x, s \in \mathbb{R}^n$ let $x^s = \prod_{i=1}^n x_i^{s_i}$. Rate of reaction $y_j \rightarrow y'_j$ is given by,

$$v_j(x) = k_{y_j \rightarrow y'_j} x^y.$$ 

System of ODEs associated with mass-action system $(\mathcal{S}, \mathcal{C}, \mathcal{R}, k)$ is given by

$$\dot{x} = Nv(x, k), \quad x \in D \subset \mathbb{R}^n \quad (2)$$

where $v(x, k) = [v_1(x), \ldots, v_m(x)]^T$ and $N = [y'_1 - y_1, \ldots, y'_m - y_m]$ is an $\mathbb{R}^n \times \mathbb{R}^m$ called the stoichiometric matrix.

Local existence and uniqueness of the solutions to the initial value problem (with condition that $x(0) \geq 0$) are guaranteed [61] and solutions will always stay positive, namely $x(t) \geq 0$ ($x(t) > 0$) if $x(0) \geq 0$ ($x(0) > 0$) [60]. From now on, all solutions discussed in the paper will be assumed to be positive, i.e., $x \in \mathbb{R}_n^+$. Important dynamic behaviour is linked to the Jacobian matrix evaluated at some positive $x$, $(J(x))$:

1. presence of a 0 eigenvalue of $J(x)$ is a necessary condition for multiple steady states.

2. presence of a pair of imaginary eigenvalues $\pm i\omega$ of $J(x)$ is a necessary condition for presence of periodic solutions.

3. if phase space $D$ is a rectangular open domain of $\mathbb{R}^n$ and $J(x)$ is a weakly $P$-matrix for all $x \in D$, then function $Nv(x)$ is injective (Gale and Nikaidô [23]).

Any matrix $A$ is a weakly $P$-matrix if the determinant of $A$ is positive and all other principal minors are nonnegative.

For open system CRNs where all species have inflow and outflow reactions, stronger condition than (3) holds: global injectivity follows if $\text{det}(J(x)) \neq 0$ (Craciun and Feinberg [10]).

For mass-action chemical reaction systems the Jacobian matrix (2) evaluated at some $x \in \mathbb{R}_n^+$ has specific form. Entries are of the form,

$$J_{ik} = \sum_{j=1}^m (\beta_{ji} - \alpha_{ji})\alpha_{jk} \frac{v_j(x)}{x_k} \quad (3)$$

for any $i, k \in \{1, \ldots, n\}$. Note that each Jacobian entry quantifies the influence that a species $A_k$ exerts on another species $A_i$. Each entry can be decomposed into the influence each species has on the other via the $j$-th reaction,

$$J_{ijk}(x) = (\beta_{ji} - \alpha_{ji})\alpha_{jk} \frac{v_j(x)}{x_k}. \quad (4)$$

Matrix properties, such as the $P$-matrix property, are determined by the matrix structure. All three graphs that will be introduced encode information about the Jacobian matrix and their topologies are related to particular decompositions of the Jacobian matrix.

The interaction graph edges are in one-to-one correspondence with the entries of the Jacobian matrix, while directed bipartite graph and SR graph rely on the
specific decomposition of the entry into terms of type \( J_{ijk} \). As will be shown, the interaction graph has an edge between vertices \( A_k \) and \( A_i \) if \( J_{ki} \neq 0 \). SR and DB graphs have two edges between species \( A_k \) and \( A_i \) linked by reaction vertex \( y_j \rightarrow y'_j \) if \( J_{kji} \neq 0 \).

The Jacobian matrix evaluated at any positive steady state \( x \) can be written in the vector form as

\[
J(x) = N \text{diag}(v(x)) \kappa^T \text{diag}(1/x)
\]

where \( \text{diag} \) denotes a diagonal matrix, \( \text{diag}(1/x) = \text{diag}(1/x_1, \ldots, 1/x_n) \) and \( \kappa \) is a kinetic matrix \( \kappa = [y_1, \ldots, y_m] \). The SR and DB graphs cycles are in correspondence with the loops of the square submatrices of the stoichiometic matrix \( N \) and the kinetic matrix \( \kappa \). An interested reader should refer to [11] for details of the correspondence.

2.1 Example model

In the next few sections it will be described how graph theoretic approaches encode the information from the reaction network and hence provide easy but also diverse ways for studying dynamical CRNs. The example considered throughout this paper is a well-known model describing activity of the mitogen-activated protein kinase (MAPK) [39, 9]. MAPKs play an important part in the signalling processes of eukaryotic cells. The model is known to exhibit bistability, namely it has up to two stable equilibria, when parameters are chosen appropriately. In this model we denote the MAPK kinase by \( S_0 \), and the molecules with single and double phosphorylation \( S_1 \) and \( S_2 \), respectively. The symbols \( E \) and \( F \) represent MEK and the phosphatase. The reaction scheme for the model is:

\[
\begin{align*}
E + S_0 & \rightleftharpoons ES_0 \rightarrow E + S_1 \rightleftharpoons ES_1 \rightarrow E + S_2 \\
F + S_2 & \rightleftharpoons FS_2 \rightarrow F + S_1 \rightleftharpoons FS_1 \rightarrow F + S_0.
\end{align*}
\]

Label the concentrations of species \( E, F, S_0, S_1, S_2, ES_0, ES_1, FS_1 \) and \( FS_2 \), from \( x_1 \) up to \( x_9 \), respectively. Let \( C_1 \) to \( C_{10} \) denote the complexes \( E + S_0 \), \( ES_0 \), \( E + S_1 \), \( S_1 \), \( ES_1 \), \( ES_2 \), \( F + S_2 \), \( FS_1 \) and \( F + S_0 \). With these definitions, the reaction system is associated with the following set of ODEs:
The interaction graph be analysed via interaction graphs. with other kinetics (where rate reactions are not necessarily monomials) can also choice of mass-action kinetics is a modeling assumption, chemical reaction networks and hence its application is also widespread in biology, where models are often not exerts on another. The interaction graph can be constructed for all classes of ODEs The interaction graph conveys the information about the influence that one species 3 Interaction graph

A edge from \( A \) to \( J \) (resp. \( G \)) example of one such system. A cycle in the graph

There are important classes of systems for which the sign pattern of the Jacobian matrix doesn’t change with change in variables and so structure of interaction graphs also doesn’t depend on species concentrations. The MAPK model in Figure 1 is an example of one such system. A cycle in the graph \( G_{\text{int}}(x) \) is a sequence of distinct vertices \( A_{i_1}, \ldots, A_{i_k} \) such that there is an edge connecting \( A_{i_1} \) to \( A_{i_2} \), \( A_{i_2} \) to \( A_{i_3} \) and so on finishing with an edge between \( A_{i_k} \) and \( A_{i_1} \). The length of the cycle is the number of vertices that it contains. Special case of a cycle containing only one vertex is called a loop. A loop has length one.

Each cycle is endowed with a sign, which is the product of the signs of its edges. In Figure 1 vertices \( S_1 \), \( FS_1 \) and \( F \) form a cycle. This cycle is has a negative sign.

\[
\begin{align*}
\dot{x}_1 &= -k_{C_1-C_2}x_1x_3 + (k_{C_2-C_1} + k_{C_3-C_4})x_6 - k_{C_5-C_4}x_1x_4 + (k_{C_1-C_3} + k_{C_4-C_5})x_7 \\
\dot{x}_2 &= -k_{C_6-C_5}x_2x_4 + (k_{C_7-C_6} + k_{C_1-C_9})x_9 - k_{C_8-C_9}x_2x_5 + (k_{C_9-C_8} + k_{C_3-C_10})x_8 \\
\dot{x}_3 &= -k_{C_1-C_2}u_1x_3 + k_{C_2-C_1}x_6 + k_{C_3-C_10}x_9 \\
\dot{x}_4 &= k_{C_2-C_3}x_6 - k_{C_3-C_2}x_1x_4 + k_{C_4-C_3}x_7 + k_{C_7-C_3}x_8 - k_{C_8-C_9}x_2x_5 + k_{C_9-C_8}x_9 \\
\dot{x}_5 &= k_{C_4-C_5}x_7 - k_{C_6-C_7}x_2x_4 + k_{C_7-C_4}x_8 \\
\dot{x}_6 &= k_{C_1-C_2}x_1x_3 - (k_{C_2-C_1} + k_{C_2-C_3})x_6 \\
\dot{x}_7 &= k_{C_3-C_4}x_1x_4 - (k_{C_4-C_3} + k_{C_4-C_5})x_7 \\
\dot{x}_8 &= k_{C_6-C_7}x_2x_5 - (k_{C_9-C_8} + k_{C_3-C_10})x_8 \\
\dot{x}_9 &= k_{C_6-C_7}x_2x_4 - (k_{C_7-C_6} + k_{C_7-C_3})x_9.
\end{align*}
\]

In the following sections MAPK model will be analysed by different graph theoretic approaches.

3 Interaction graph

The interaction graph conveys the information about the influence that one species exerts on another. The interaction graph can be constructed for all classes of ODEs and hence its application is also widespread in biology, where models are often not of polynomial structure. This versatility is also useful in chemical models. Since choice of mass-action kinetics is a modeling assumption, chemical reaction networks with other kinetics (where rate reactions are not necessarily monomials) can also be analysed via interaction graphs.

**Definition 3.1.** The interaction graph \( G_{\text{int}}(x) \) of CRN at a point \( x \in \mathbb{R}^n \) is a finite oriented graph with species \( V = \{A_1, \ldots, A_n\} \) as set of vertices and with a directed edge from \( A_i \) to \( A_k \) if \( J_{ik}(x) \neq 0 \). Edge is positive (resp. negative) if \( J_{ik}(x) > 0 \) (resp. \( J_{ik}(x) < 0 \)).

Interaction graph of MAPK cascade is presented in Figure 1. The structure of the interaction graph depends on the species concentrations, since edges can change sign for different species concentrations \( x \) and in the process can also vanish when there is no interaction, namely when \( J_{ik} = 0 \).

There are important classes of systems for which the sign pattern of the Jacobian matrix doesn’t change with change in variables and so structure of interaction graphs also doesn’t depend on species concentrations. The MAPK model in Figure 1 is an example of one such system. A cycle in the graph \( G_{\text{int}}(x) \) is a sequence of distinct vertices \( A_{i_1}, \ldots, A_{i_k} \) such that there is an edge connecting \( A_{i_1} \) to \( A_{i_2} \), \( A_{i_2} \) to \( A_{i_3} \) and so on finishing with an edge between \( A_{i_k} \) and \( A_{i_1} \). The length of the cycle is the number of vertices that it contains. Special case of a cycle containing only one vertex is called a loop. A loop has length one.

Each cycle is endowed with a sign, which is the product of the signs of its edges. In Figure 1 vertices \( S_1 \), \( FS_1 \) and \( F \) form a cycle. This cycle is has a negative sign.
Two cycles are called disjoint, if they don’t share any common vertices. A union of disjoint cycles which contains all vertices of $G_{int}(x)$ is called a nucleus. In some literature it is referred to as a Hamiltonian hooping [53]. Nucleus is also endowed with a sign $(-1)^{p+1}$, where $p$ is the number of positive cycles of the nucleus. A nucleus (or a cycle) is variable if at least one of its edges changes sign for different values of $x$. A nucleus (or cycle) is called ambiguous when its sign varies with $x$.

3.1 Results

Determinant and minors of the Jacobian matrix can be determined via Leibnitz formula and calculated over permutations of rows and columns of the Jacobian (or its square submatrix, in the case of minors). Since edges of the interaction graph are in one-to-one correspondence with the nonzero entries of the Jacobian matrix, a set of entries of the Jacobian whose rows and columns are in cyclic permutation are also in one-to-one correspondence with the cycles of the interaction graph. Hence, it is clear that cycles of the interaction graph are strongly linked to the system dynamics.

In early 1980s René Thomas [54] conjectured that any dynamical system with multiple steady states must have a positive cycle while a system with stable oscillations must contain at least one negative cycle (of length at least two). These conjectures are very intuitive, as we will now demonstrate on the interaction graph of the MAPK cascade. Consider the negative cycle of species $S_1$, $F$ and $FS_1$. Increasing the concentration of $S_1$ will increase the concentration of $FS_1$ which, as it
unbinds, will in turn increase the concentration of $F$. But with higher concentration of $F$, concentration of $S_1$ will decrease. So an initial increase in concentration of $F$ along the negative cycle leads to a decrease in its concentration, and this explains how negative cycles could possibly promote oscillatory behaviour. Now take the positive cycle between species $ES_1$, $S_2$, $S_2$ and $S_1$. An increase in concentration of $S_1$ will lead to a further increase in its concentration around the cycle. Perturbation of $S_1$ pushes the system to a new basin of attraction and possibly a new equilibrium. Thomas’s first conjecture was proven by Soulé in [53]:

**Theorem 3.2 (Thomas-Soulé).** If a system has no positive cycles in $G_{int}(x)$ for any $x$, then it cannot have two or more stable equilibria present simultaneously.

Aside from the relationship between cycles and Jacobian entries, which lead to conditions about positivity of the determinant of $-J$ and its minors, Soulé made use of Gale-Nikaidō theory [23] to show that the reaction system is injective and hence cannot exhibit multistationarity.

Thomas’ second conjecture has been shown to be correct for quasimonotone systems by Snoussi [52] and Gouzé [28]. Quasimonotone systems are systems whose interaction graph edges are not sign-changeable. The MAPK cascade is a quasimonotone system and it has at least one positive cycle, from Figure 1, hence it may have multiple steady states.

Aside from Thomas’ conjecture, the negative cycle is linked to various dynamics. For a steady state $x$, a negative cycle in the graph $G_{int}(x)$ is necessary in order for a steady state to be stable [56, 47]. Moreover, a negative cycle of length one is necessary for the presence of general attractor [56], a compact subset of the phase space which is invariant under flow and which attracts a fundamental family of open neighbourhoods. Both of these types of cycles are present in the interaction graph for the MAPK cascade.

To understand dynamics of the system one must understand how a particular cycle fits into the full network landscape. Consider the positive cycle of species $ES_1$, $S_2$, $S_2$ and $S_1$. It was previously concluded that when this cycle is isolated, an increase in $S_1$ will lead to further increases of $S_1$ if only interactions along the positive cycle are considered. However, influence of species outside the cycle might give a different result. Consider again the negative cycle of species $S_1$, $FS_1$ and $F$. Clearly increasing $S_1$ will decrease its concentration if we follow negative cycle and this might offset the increase from the positive cycle. Concentration of $S_1$ could end up being smaller than the initial concentration.

Conjectures about the dynamics and their relation to the full network structure have also been explored. René Thomas [55] posited that the appearance of both negative and positive cycles is necessary for chaotic dynamics. This conjecture was discussed in detail by Toni in [58]. On the full scale of the system Kaufmann and Thomas [57] conjectured that presence of a variable nucleus or presence of two nuclei of opposite sign are necessary for a system to display multistationarity. Recently together with Soulé [56] they proved their conjecture for a class of systems which encompass mass action chemical reaction systems.

In the MAPK case in Figure 1 least two nuclei with opposite sign can be identified. First nuclei is composed of a union of cycles $\{E, ES_0, S_0\}$, $\{F, FS_1\}$ and $\{S_1, ES_1, S_2, FS_2\}$. The second has cycles $\{FS_1, S_1, F\}$, $\{S_2, FS_2\}$, $\{ES_1, E\}$ and $\{S_0, ES_0\}$. 
4 SR graph

The species-reaction graph was introduced in a paper of Craciun and Feinberg in [11] and it is a special type of the Petri-net graph, commonly used in computer science.

**Definition 4.1.** Species-reaction graph $G_{SR}$ is an unoriented bipartite graph with two set of vertices: of species $\{A_1, \ldots, A_n\}$ and of reactions, $\{y_1 \to y'_1, \ldots, y_m \to y'_m\}$ which exclude inflow and outflow reactions. If $\alpha_{ji} > 0$ then there is an edge (labeled $y_j$) between species vertex $A_i$ and reaction vertex $y_j \to y'_j$. If $\beta_{ji} > 0$ then there is an edge (labeled $y'_j$) between $A_i$ and $y_j \to y'_j$.

Edges that are connected by same reaction node and have identical complex label are called a c-pair. Aside from the complex label, each edge in the SR graph also has an attached stoichiometric coefficient.

An edge between species $A_i$ and reaction $y_j \to y'_j$ can be labeled by $(A_i, y_j \to y'_j)$ or $(y_j \to y'_j, A_i)$. A path in an SR graph is a collection of distinct edges where first vertex of an edge is the last vertex of the preceding edge, for example, $(A_{i_1}, y_{i_2} \to y'_{i_2}, A_{i_2}), (A_{i_2}, y_{i_3} \to y'_{i_3})$. A cycle is a path where first vertex of the first edge (of the path) is the last vertex of the last edge in the path. For example, $(A_{i_1}, y_{i_2} \to y'_{i_2}), (y_{i_2} \to y'_{i_2}, A_{i_2}), (A_{i_2}, y_{i_3} \to y'_{i_3}), \ldots, (y_{i_k} \to y'_{i_k}, A_{i_1})$ is a $k$ order cycle.

Cycles of SR graphs are separated into classes depending on the number of c-pairs that they contain. A cycle is an $e$ cycle if it has an even number of c-pairs, otherwise it is an $o$-cycle.

Another class is also important. A cycle is called an $s$-cycle, if when the stoichiometric coefficient labels of the edges are alternately multiplied and divided, the final product is 1. A simple path from a species vertex to a reaction vertex in $G_{SR}$ is defined to be an S-to-R chain. Two cycles in $G_{SR}$ are said to have an S-to-R intersection if their common edges belong to a S-to-R chain or a disjoint union of several chains.

The SR graph also has an oriented version, called the OSR graph. All theorems presented by Craciun and Feinberg are related to the properties of the SR graph and only use the oriented version in the proofs. Hence, for the sake of brevity, we will not overview their construction. Interested reader should consult [11].

In the interaction graph, we mentioned that every cycle presents a set of nonzero coefficients of the Jacobian whose row and column are in cyclic permutation. Contribution of Craciun and Feinberg theory is that every entry $J_{ik}$ is split into its component terms $J_{ijk}$ related to the contribution each $j$-th reaction rate makes to the entry $J_{ik}$. This split allows for finer results relating the structure of the graph to the values of the minors of the Jacobian. In contrast to the interaction graph, each path between two species vertices encodes information about the reaction which links the two species. For example, look at the cycle between species $E$ and $S_1$ in the interaction graph of MAPK network in Figure 1. The interaction between these two species depends on two reactions.

However, there are some important differences. Cycles in the interaction graph are in direct correspondence with the cycles in the entries of the Jacobian matrix. Cycles in SR graph are related to the cycles in the stoichiometric matrix and kinetic matrix, described in Lemma 5.1 and 5.2 in [11]. Cycles in the interaction graph and

\[\text{some reactions could be reversible}\]
Figure 2: SR graph of the MAPK chemical reaction network.
SR graph do not have a one-to-one correspondence. For example, take any one of the cycles in the interaction graph of the MAPK network in Figure 1 with species \( E, ES_0 \) and \( S_0 \). In the SR graph in Figure 2 these three species are not in a cycle. In order for a cycle to exist in the SR graph, there would have to be a third reaction vertex adjacent to two edges that are incident to species \( S_0 \) and \( ES_0 \).

An advantage of Craciun-Feinberg theory is that the theory is aimed at mass-action kinetic systems and this specific structure can lead to stronger results.

**Theorem 4.2.** Consider some open CRN where each species takes part in an inflow and an outflow reaction. If its associated SR graph \( G_{SR} \) contains only cycles that are \( o \)-cycles or \( s \)-cycles, and such that no two \( e \)-cycles have an \( S \)-to-\( R \) intersection, then the reaction network is injective.

The proof presented by Craciun and Feinberg in [11] relies on two parts: that cycles conditions imply that the Jacobian matrix is nonsingular and then that this result as well as the polynomial structure of the ODEs (due to mass-action kinetics and inflow/outflow reactions) imply that the vector field is injective.

### 4.1 Directed bipartite graph

The directed bipartite graph was developed by Ivanova [33, 34, 59] and more rigorously analysed by Mincheva and Roussel in [41, 42]. It is very closely related to the species-reaction graph. The graph has been applied to some real mechanisms by Ermakov, Goldstein et al. [16, 24, 26, 27] and analysed for small networks with up to four species [15, 17, 25].

To write the system as a directed bipartite graph, all reversible reactions have to be rewritten as pairs of irreversible reactions. For example,

\[
E + S_0 \rightleftharpoons ES_0
\]

becomes two reactions:

\[
E + S_0 \rightarrow ES_0 \quad \text{and} \quad E + S_0 \leftarrow ES_0.
\]

**Definition 4.3.** A directed bipartite graph associated with a CRN is a graph \( G_{DB} \) with set of vertices partitioned into a set of species \( \{A_1, \ldots, A_n\} \) and set of reactions \( \{y_1 \rightarrow y'_1, \ldots, y_m \rightarrow y'_m\} \). If \( \alpha_{jk} > 0 \) there exists a directed edge from \( A_k \) to reaction vertex \( y_j \rightarrow y'_j \). Similarly, if \( \beta_{ji} > 0 \) there exists a directed edge from \( y_j \rightarrow y'_j \) to \( A_k \).

The DB graph of MAPK is shown in Figure 3. It has more reaction vertices than the SR graph. A path is a union of two edges \( (A_k, y_j \rightarrow y'_j) \) and \( (y_j \rightarrow y'_j, A_i) \) that share a common reaction vertex. A path is positive if \( \alpha_{jk} \beta_{ji} > 0 \). Positive path is denoted by \( [A_k, y_j \rightarrow y'_j, A_i] \). A path is negative if \( \alpha_{jk} \beta_{ji} > 0 \). A negative path is denoted by \( [A_k, y_j \rightarrow y'_j, A_i] \). If a path is negative, then \( (A_i, y_j \rightarrow y'_j) \) and \( (y_j \rightarrow y'_j, A_k) \) also form a different negative path, \( [A_i, y_j \rightarrow y'_j, A_k] \), since \( \alpha_{ji} \alpha_{ki} > 0 \). Cycles are defined differently in the directed bipartite graph than in the SR graph. A cycle in the directed bipartite graph is a union of distinct paths \( (A_{i_1}, y_{i_1} \rightarrow y'_{i_1}, A_{i_2}), (A_{i_2}, y_{i_2} \rightarrow y'_{i_2}, A_{i_3}) \) until \( (A_{i_k}, y_{i_k} \rightarrow y'_{i_k}, A_{i_{k+1}}) \). A \( c \)-pair is considered to be a cycle, because path from one species vertex \( A_1 \) to another \( A_2 \) is considered to be different from a path from \( A_2 \) to \( A_1 \), and the two together form a cycle. For example, a pair
Figure 3: DB graph of the MAPK network.
of edges joining \( S_1 \) and \( F \) via reaction vertex \( F + S_1 \rightarrow FS_1 \) form a cycle of two negative paths.

Every cycle in the interaction graph corresponds to one or more cycles in the directed bipartite graph. This occurs because each edge in interaction graph represents an interaction between two species, say \( S_i \) and \( S_j \), so namely, \( J_{ji} \), and this Jacobian entry can be decomposed by type of reaction, which in directed bipartite graph is represented by several paths, each with a different reaction vertex.

The main advantage of this graphical interpretation is the relation between each coefficient of the expansions of the determinant of \(-J\) (or minor of \(J\)) is explicitly stated in terms of the edges and cycles of \(G_{DB}\).

A subgraph \( g \) of \( G_{DB} \) is a union of edges and cycles which are disjoint. The order of a subgraph is defined to be the number of species vertices contained in the subgraph. The set of all subgraphs \( g \) of order \( k \) that share the same set of vertices is called a fragment of order \( k \) and denoted by \( S_k(\{i_1, \ldots, i_k\}) \) where \( i_1, \ldots, i_k \) are the indices of the species and \( j_1, \ldots, j_k \) are index subset of the reactions in the vertex sets.

With these structures the graph-theoretic formula for the coefficients \( a_k \) of characteristic polynomial \((\lambda I - J)\) and the determinant of \(-J\) (which is \(a_n\)) can be derived in terms of fragments. If \( \{g\} \) are a set of all subgraphs of \( G_{DB} \) of order \( n \), then

\[
det(-J) = \sum_{g \subseteq G} K_g x_{i_1} \cdots x_{i_k} \tag{5}
\]

where each coefficient \( K_g \) is a product of stoichiometric labels over all edges and all \( t_g \) cycles \( (C) \) of \( g \)

\[
K_g = (-1)^{t_g} \prod_{[A_k, B_j] \in g} \alpha_{j_k}^2 \prod_{C \in g} K_C \tag{6}
\]

and \( K_C \) is a product of all stoichiometric labels over all paths (negative/positive) of each cycle \( C \in g \)

\[
K_C = \prod_{[A_k, B_j, A_l] \in C} \frac{(-\alpha_{j_k} \alpha_{j_l})}{(-\alpha_{j_k} \alpha_{j_l})} \prod_{[A_k, B_j, A_l] \in C} \alpha_{j_k}^{\beta_{jl}}. \tag{6}
\]

Likewise, one can derive determinant formulas for the minors of the Jacobian matrix. Each coefficient \( a_k \) of the characteristic polynomial,

\[
det(\lambda I - J) = \lambda^n + a_1\lambda^{n-1} + \ldots + a_n. \tag{7}
\]

is the sum of all principal minors \( M(-J)(i_1, \ldots, i_k) \) of order \( k \), where \( 1 \leq i_1 < \ldots < i_k \leq n \). If \( \text{rank}(N) = r < n \), then the last \((n-r)\) coefficients, \( a_{n-r} = \ldots = a_n = 0 \).

Similar to the previous formula,

\[
a_k = \sum_{S_k(\{i_1, \ldots, i_k\})} K_S \frac{\prod_{j_1} \cdots \prod_{j_k} v_{j_i}}{x_{i_1} \cdots x_{i_k}} \tag{8}
\]

where contribution of fragment \( S_k \) is,

\[
K_{S_k} = \sum_{g \subseteq S_k(\{i_1, \ldots, i_k\})} K_g. \tag{8}
\]
Derivation of both formulas is stated in [41]. A fragment $S_k$ such that $K_{sk} < 0$ is called a critical fragment. It contains at least one subgraph with an odd number of positive cycles, namely $K_y < 0$.

In fact, from the formula above, each subgraph which is the disjoint union of cycles and edges just contain the same number of species and reaction vertices. Hence, the cycles formed with repeating reaction vertices do not appear in the calculation of the coefficients of the determinant of Jacobian or any of the characteristic polynomial coefficients.

Despite the definition of the cycles, these notions are not used in the work of Mincheva and Roussel [41] and instead all attention is given to the critical fragments.

Critical fragment is a useful concept for manipulation of the characteristic polynomial, $det(\lambda I - J)$.

Since critical fragments are defined to be negative coefficients of the expansion of the determinant of $-J$ (or the minors) necessary conditions for any sort of instability reduce to a search for critical fragments [41, 42].

Directed bipartite graph has some advantages and disadvantages over the SR graph. A comparison of the conditions for spatial instability (such as Turing instability) that arise from DB graph and SR graph have been explored in [14]. For delay-induced instability, described in [43], the directed bipartite graph seems more appropriate than the SR graph.

5 Conclusion

Graph theoretic approaches to studying chemical reaction networks have been steadily developing in the past few decades. Graphs are useful tools for studying chemical reaction networks, because they can deal with issues at which some other methods can fail. For example, network size and lack of information about the parameter constants. Here three graphs: interaction graph, species-reaction graph and directed bipartite graph have been described. Comparisons between have been drawn, where this was possible. Further development is necessary if one wishes to gain a better understanding of the underlying mechanisms.

Acknowledgements

The author was supported by the Commonwealth Doctoral Scholarship. This paper is based on review of [13].

References


[13] M. Domijan and M. Kirkilionis, Graph Theory and Qualitative Analysis of Reaction Networks, accepted to Networks and Heterogeneous Media.


[34] A. N. Ivanova and B. L. Tarnopolskii, *One approach of the determination of a number of qualitative features in the behaviour of kinetic systems, and realization of this approach in a computer (critical conditions, autooscillations)*, Kinet. Katal., 20 (1979), 1541–1548.


