Contribution to:

Mathematical methods to gain biological insight

Sander Hille
Mathematical Institute
Leiden University

15th November 2015

Chapter 1

Chemical Reaction Networks

By Sander Hille

At cellular level Life's processes consist of the electro-chemical interactions between a large number of different chemical compounds: ions, small signaling molecules, peptides, proteins, RNA, DNA,... In this chapter we focus on the mathematical modeling of (large) networks of chemical reactions. Theory on structure of convex sets shall provide a unique set of so-called Elementary Flux Modes (EFMs), introduced by Schuster and coworkers [9, 10]. These can be interpreted as elementary modes of operation of the network at steady state. Moreover, these are '*pathways*': a directed path of chemical reaction starting from a particular set of substrates leading to a particular set of products.

Although the concept of EFM is applicable to any chemical reaction network, it has been developed for and is particularly useful in the study of *metabolic networks*. Metabolism is the collective of chemical reactions in an organism that sustains life, i.e. realizes the break down of nutrients into a limited set of compounds that is subsequently used to assemble more complicated chemical building blocks, specific to the organism. For example, in the metabolic process of glycolysis one molecule of glucose is step-wise broken down by various enzymes into two molecules of pyruvate, harvesting 'energy' on the way that is stored in the form of two molecules of adenosine-tri-phosphate (ATP). Pyruvate and ATP are subsequently used in various other metabolic processes. In a fixed environment, metabolic processes quickly settle to a steady state, typically. In contrast, the function of *signaling networks* in organisms derives primarily from temporal changes in behaviour of the network. Therefore, other techniques than EFM-analysis are more appropriate to study these.

In this chapter we focus on EFM-analysis for metabolic networks. Central in the modeling is the *stoichiometry* of the reactions: the precise number of molecules of specific species that are involved in a single step of a reaction in the network. It is the stoichiometry of the network that limits the *possible states* of the network at steady state, independent of the specific kinetics of individual reactions involved. This type of analysis is known as *constrained-based stoichiometric analysis* of the network. The *specific steady state* of a network does depend on the specific kinetics and parameter settings of each reaction.

1.1 The stoichiometry of a reaction network

A reaction network is specified by a collection $\mathcal{M} := \{\mathrm{M}_1, \ldots, \mathrm{M}_m\}$ of chemical compounds (molecules, called '*metabolites*' in a metabolic network) that are distinguished either by their chemical structure, spatial configuration or their physical location. The same molecule located in different physically separated compartments is modeled as different M_i. When modeling cellular metabolism a typical compartmentation is provided by the natural distinction between the environment, or exterior, of the cell in which there are nutrients of interest, like glucose, and the cytoplasm inside the cell. In bacteria the cytoplasm is the sole interior compartment. However, cells of eukaryotes like yeast, the amoeba *Dictyostelium discoideum*, animal cells or plant cells have multiple interior compartments in which parts of metabolism take place. Thus, for example glucose in the environment and glucose in the cytoplasm of a bacterium is modelled by M_i and M_j with $i \neq j$.

Species in $\mathcal{M}$ that reside in the same compartment can be transformed into species in the very same compartment by means of *internal reactions*, labelled $\mathrm{R}_1, \ldots, \mathrm{R}_n$. Species in a compartment can be transported and/or transformed into species in a neighbouring compartment by *exchange reactions*. These are labelled $\mathrm{ER}_1, \ldots, \mathrm{ER}_e$.

$$\mathcal{R} := \{\mathrm{R}_1, \ldots, \mathrm{R}_n\} \cup \{\mathrm{ER}_1, \ldots, \mathrm{ER}_e\}$$

is the total set of reactions in the network. Internal and exchange reactions are distinguished because the latter require accounting for possibly different volumes of compartments in which substrates and products reside (see below). Reactions are considered to be either (essentially) *irreversible* or *reversible*. Let $\mathrm{Irr}(\mathcal{R})$ and $\mathrm{Rev}(\mathcal{R})$ denote these disjoint subsets. A characteristic feature of metabolic networks is that almost every reaction is catalyzed by a specific enzyme. When this enzyme is not present the corresponding reaction is highly improbable to occur. Effectively, it does not.

The *stoichiometric matrix* S for the reaction network is an $m \times (n+e)$-matrix with integer entries that codes how molecular numbers of each species in $\mathcal{M}$ are changed when any reaction from $\mathcal{R}$ is executed once, in positive direction in case of reversible reactions. In this matrix a negative integer indicates the number of a particular metabolite that is used as substrate in a single reaction step (in positive direction), while positive integers indicate the number of that metabolite being produced per step. Accordingly, the disjoint partitioning of $\mathcal{R}$ into $\mathrm{Irr}(\mathcal{R})$ and $\mathrm{Rev}(\mathcal{R})$ together with the specification of S defines for each reversible reaction a positive direction.

Definition 1.1.1 *A* **chemical reaction network** $\mathcal{N}$ *consists of the specification of* $\mathcal{M}$, $\mathrm{Irr}(\mathcal{R})$ *and* $\mathrm{Rev}(\mathcal{R})$ *(hence* $\mathcal{R}$*) together with the stoichiometric matrix* S*.*

Figure 1.1 provides an example of an (artificial) chemical reaction network and its associated stoichiometric matrix and a graph representation.

R1: $M_1 + M_2 \rightarrow M_3$

R2: $M_3 \rightarrow M_4 + 2M_5$

R3: $M_4 \rightleftarrows M_2$ *(reversible)*

Stoichiometric matrix:

$$S = \begin{array}{c} \\ M_1 \\ M_2 \\ M_3 \\ M_4 \\ M_5 \end{array} \begin{array}{c} \begin{array}{cccc} R_1 & R_2 & R_3 & \quad ER_1 \quad ER_2 \end{array} \\ \left[\begin{array}{cccc:cc} -1 & 0 & 0 & 1 & 0 \\ -1 & 0 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 \\ 0 & 2 & 0 & 0 & -1 \end{array}\right] \end{array}$$

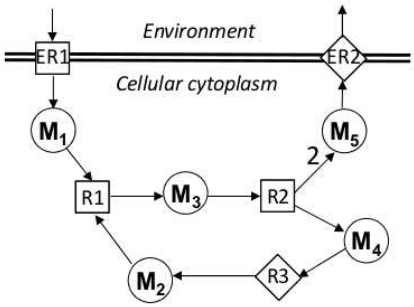

Figure 1.1: *A small reaction network with three internal reactions (R1, R2 and R3) and two exchange reactions (ER1 and ER2). ER2 and R3 are reversible, all other reactions are irreversible. The positive direction of the reactions is indicated by the arrows. The corresponding stoichiometric matrix is presented.*

1.2 The fundamental master equation

We formulate a mathematical model for the dynamics of a chemical reaction network that is valid under the assumptions that (1) each compartment is well-stirred, i.e. all molecules are homogeneously distributed within, and (2) the molecular numbers are sufficiently high such that stochastic effects can be ignored. The state of the system at time t can then be characterized well by the number $n_i(t)$ of molecules of species M_i present at that time and the volumes of the compartments. Denote by $V_i(t)$ the volume at time t of the compartment in which M_i is located. Equivalently, one can use concentrations x_i and volumes V_i. However, if there are exchange reactions between compartments with different volumes one can start best with a number-description and derive from that a concentration-description to get correct equations.

So let $n := (n_1, \ldots, n_m)^T$ be the state vector of metabolite numbers. The *reaction flux* of a reaction $r \in \mathcal{R}$ at a time t, $\Phi_r(t)$ is the (average) number $\Phi_r(t)$ of reaction r that takes place per unit time at time t, with the convention that $\Phi_r(t) \geq 0$ for $r \in \mathrm{Rev}(\mathcal{R})$ when the there are more or the same number of reaction per unit time in positive direction than there are in negative direction. Otherwise it is negative. So

$$\begin{cases} \Phi_r(t) \geq 0, & \text{for all } t \geq 0,\ r \in \mathrm{Irr}(\mathcal{R}), \\ \Phi_r(t) \in \mathbb{R}, & \text{for all } t \geq 0,\ r \in \mathrm{Rev}(\mathcal{R}). \end{cases} \tag{1.1}$$

Let $\Phi = (\Phi_1, \ldots, \Phi_n, \Phi_{n+1}, \ldots \Phi_{n+e})^T$ be the *reaction flux distribution vector*, or simply the *flux distribution*. Here, the first n reaction fluxes correspond to the internal reactions, while the last e fluxes correspond to the exchange fluxes. Instead of Φ_r one can use *reaction flux density* v_r, which is the (average) number of reactions r that take place per unit volume per unit time. v then denotes the corresponding vector of v_r, $r \in \mathcal{R}$. Notice, that the definition of reaction flux density requires a rather arbitrary decision what compartment to take for the

reference volume of an exchange reaction, since it connects two compartments and can be considered to reside in either of them, or both. For internal reactions of course one simply takes the volume of the compartment in which the takes place as reference volume to define the reaction flux density.

Each Φ_r will be a function of metabolite *concentrations* and parameter values. That is, it will depend on the $x_i = n_i/V_i$. For example, if M_j is obtained as product of an enzyme-catalyzed reaction of Michaelis-Menten type from a substrate M_i, then

$$\Phi_r = \frac{\nu_r x_i}{\kappa_r + x_i}. \tag{1.2}$$

(Recall Section [**ref to S2.1**]). More complicated enzymatic reactions lead to more complicated expressions for Φ_r as function of concentrations in x (see Section [**ref to S2.3**]). We simply write $\Phi = \Phi(x; p)$ to express the dependence of the flux distribution on the concentration vector and all parameters, now summarized in the vector p.

Then we have the *fundamental master equation*

$$\frac{dn}{dt} = S\Phi(x; p). \tag{1.3}$$

Writing $n(t) = V(t)x(t)$, where $V = \mathrm{diag}(V_1, \ldots, V_m)$ is the $m \times m$ diagonal matrix with the volumes V_i on the diagonal, one obtains

$$\frac{dx}{dt} = V^{-1}S\Phi(x; p) - V^{-1}\frac{dV}{dt}x. \tag{1.4}$$

The last term in (1.4) is called the *dilution term*. It accounts for change in concentration due to change in volume of compartments rather than change in molecular numbers. The diagonal matrix $V^{-1}\frac{dV}{dt}$ consists of the relative rates of change of the compartments. Often, the dilution term is ignored, e.g. when volumes hardly change or when the first (reaction) term in (1.4) is dominant.

In the literature, one most often encounters situations in which nutrients in the environment are steadily supplied. It is an *open system*. These nutrients are not included as metabolites in $\mathcal{M}$ then, because their number in the environment is kept constant due to active supply or great abundance. Moreover, only a single internal compartment is considered (i.e. the cytoplasm in cells). Then V is a multiple of the identity matrix. Ignoring dilution, (1.4) now reduces to

$$\frac{dx}{dt} = V^{-1}S\Phi(x; p) \qquad \text{or} \qquad \frac{dx}{dt} = Sv(x; p). \tag{1.5}$$

Recall that v is the vector of reaction flux densities taken with reference to the volume of the single internal compartment in the model. It is the second equation in (1.5) that is therefore typically encountered as model for metabolic network dynamics, without the explicit mentioning of the various underlying assumptions.

1.3 The steady state flux cone

Analysis of the detailed dynamics of the (large) system of non-linear ordinary differential equations defined by (1.3), (1.4) or (1.5) can be complicated or even

unfeasible. The reaction kinetics may not be known precisely, or when it is (some) parameter values my be unknown. Useful information can be obtained for the steady states however. Let us *ignore dilution* from here on.

For any steady state $x^* \in \mathbb{R}^m_+$ of (1.5) one must have

$$S\Phi(x^*; p) = 0. \tag{1.6}$$

From a purely mathematical perspective it is then natural to describe the associated steady state flux distribution $\Phi^* := \Phi(x^*; p)$ in terms of basis vectors for the null space of S, $\ker(S)$. However, a basis is not unique and basis vectors will generally fail to satisfy the *positivity constraint* (1.1) that is imposed on a flux distribution for the network. Linear basis vectors thus have no biological meaning. A more appropriate representation of all possible steady states by means of particular flux distributions that do satisfy the positivity constrained is needed.

To introduce such distributions, put

$$\Gamma := \{\Phi \in \mathbb{R}^{n+e} \,:\, \Phi_r \geq 0 \text{ if } r \in \mathrm{Irr}(\mathcal{R})\}.$$

Then Φ^* is in the *steady state flux cone*

$$\Gamma^* := \{\Phi \in \Gamma \,:\, S\Phi = 0\}. \tag{1.7}$$

We say that '*reaction r is involved in* $\Phi \in \Gamma^*$' if $\Phi_r \neq 0$. The *support* of Φ is

$$\mathrm{supp}(\Phi) := \{r \in \mathcal{R} \,:\, r \text{ is involved in } \Phi\}. \tag{1.8}$$

A *flux mode* is a subset of Γ^* of the form $\{\lambda\Phi \,:\, \lambda > 0\}$ for some $0 \neq \Phi \in \Gamma^*$. Thus, two flux distributions that are in the same flux mode involve the same reactions. The ratios of the reaction fluxes of the two distributions for corresponding reactions are all equal.

Γ^* is a *convex cone* in the sense that it is a *convex set*:

$$\text{if } \Phi, \Phi' \in \Gamma^*, \text{ then } \lambda\Phi + (1-\lambda)\Phi' \in \Gamma^* \text{ for all } 0 < \lambda < 1,$$

and a cone: $\Phi \in \Gamma^*$ implies $\lambda\Phi \in \Gamma^*$ for any $\lambda \geq 0$. A convex cone is called *pointed* when it does not contain any non-trivial linear subspace. So Γ^* is pointed if and only if $\Gamma^* \cap (-\Gamma^*) = \{0\}$. This happens e.g. when all reactions are irreversible.

1.4 Generation of the steady state flux cone

One says that Γ^* *is generated by a collection* $\mathcal{E}$ *of subsets of* Γ if for every $\Phi \in \Gamma^*$ there exists $E_1, \ldots E_k \in \mathcal{E}$ and $\Phi_i \in E_i$, $i = 1, \ldots, k$, such that

$$\Phi = \sum_{i=1}^{k} \lambda_i \Phi_i, \qquad \text{with } \lambda_i \geq 0. \tag{1.9}$$

Representation (1.9) for Φ is not unique in general. It is still useful however, when the collection $\mathcal{E}$ consists of 'simple' and biologically interpretable subsets. Such sets we shall now identify.

The theory of convex sets (e.g. [7]) provides suitable candidates. First, observe that Γ^* is *polyhedral*. That is, it is a finite intersection of closed half-spaces

$$H_{a_i,\nu_i} := \{v \in \mathbb{R}^{n+e} \,:\, \nu_i \cdot v \geq a_i\}$$

with $a_i \in \mathbb{R}$ and $\nu_i \in \mathbb{R}^{n+e}$. Here $\nu_i \cdot v$ denotes the standard inner product between ν_i and v. The *faces* of Γ^* are

$$\Gamma^* \cap \partial H_{a_i,\nu_i} = \Gamma^* \cap \{v \in \mathbb{R}^{n+e} \,:\, \nu_i \cdot v = a_i\}$$

and intersections of these sets. The one-dimensional faces are called *extreme rays*.

From the theory of convex sets (e.g. [7], Theorem 18.5) one obtains

Theorem 1.4.1 *Assume that Γ^* does not contain any linear subspaces, i.e. Γ^* is pointed, then Γ^* is generated by the collection of its extreme rays.*

Notice that the extreme rays are uniquely determined by the cone. However, not every steady state flux cone of a network need to be pointed. In general the set of extreme rays does not generate the full cone.

We continue by discussing three collections of subsets that do generate the steady state flux cone of a network. These are the collections of Extreme Currents, Elementary Flux Modes and Extreme Pathways.

1.4.1 Extreme Currents

Extereme Currents were defined by Clarke [3] for networks that consist of irreversible reactions only. In that case, $\Gamma^* \subset \mathbb{R}^{n+e}_+$, so Γ^* is pointed. An *Extreme Current* (EC) is a flux mode that is contained in an extreme ray of the steady state flux cone Γ^* of such a network (it is the extreme ray with 0 removed). For such a network, there are finitely many ECs and they are uniquely determined by S. In view of Theorem 1.4.1, Γ^* in this case is generated by the collection of ECs. There are algorithms to compute ECs from the stoichiometric matrix S.

1.4.2 Elementary Flux Modes

Schuster and coworksers (e.g. [9, 10]) started to consider reaction networks with reversible reaction. They introduced the idea of *simplicity*: a flux distribution $\Phi \in \Gamma^*$ is *simple* if there does not exist $\Phi' \in \Gamma^*$, $\Phi' \neq 0$, such that $\mathrm{supp}(\Phi')$ is properly contained in $\mathrm{supp}(\Phi)$. This concept is equivalent to *indecomposability*: the flux distribution Φ is indecomposable if there do not exist $\Phi_1, \Phi_2 \in \Gamma^*$, $\Phi_i \neq 0$ such that

$$\Phi = \lambda_1 \Phi_1 + \lambda_2 \Phi_2 \quad \text{for some } \lambda_1, \lambda_2 > 0 \tag{1.10}$$

and $\mathrm{supp}(\Phi_1)$ and $\mathrm{supp}(\Phi_2)$ are different proper subsets of $\mathrm{supp}(\Phi)$.

Exercise 1.4.2 *Prove that a flux distribution is simple if and only if it is indecomposable.*

From a biological point of view, simplicity or indecomposability means that the flux distribution consists of a minimal collection of reactions that together can function in metabolic steady state. If one of the reactions involved is inhibited, for example by genetic modification such that the enzyme(s) that catalyses this reaction cannot be produced – that is, the corresponding gene is 'knocked out' – then there is no subset of the reactions involved in the flux distribution that still gives a functioning metabolic steady state. Such distributions could be viewed as corresponding to *'metabolic pathways'*.

Following Schuster *et al.* [9, 10],

Definition 1.4.3 *An* Elementary Flux Mode *(EFM) is the flux mode generated by a simple or indecomposable flux distribution.*

One can show that in a reaction network with irreversible reactions only, each Extreme Current is an Elementary Flux Mode.

In fact, it will turn out, that the EFMs of a reaction network $\mathcal{N}$ in which there are reversible reactions are closely related to ECs of a reaction network $\mathcal{N}'$ with irreversible reactions only that is naturally associated to $\mathcal{N}$. $\mathcal{N}'$ is called the '*reconfigured reaction network'* associated to $\mathcal{N}$. $\mathcal{N}'$ and $\mathcal{N}$ have the same set of metabolites and irreversible reactions. Each reversible reactions $r \in \mathrm{Irr}(\mathcal{N})$ in $\mathcal{N}$ is replaced in $\mathcal{N}'$ by a pair of irreversible reactions, labelled $(r,+)$ and $(r,-)$, where $(r,+)$ corresponds to the direction of positive flux in reaction r on $\mathcal{N}$. The set $\{(r,+),(r,-)\}$ is a *reversible reaction pair*. So

$$\mathcal{R}(\mathcal{N}') = \mathrm{Irr}(\mathcal{N}') = \mathrm{Irr}(\mathcal{N}) \cup \mathrm{Rev}(\mathcal{N}) \times \{\pm\}. \tag{1.11}$$

The flux cone Γ' for $\mathcal{N}'$ is modified accordingly:

$$\Gamma' := \mathbb{R}_+^{\mathrm{Rev}(\mathcal{N}')} = \mathbb{R}_+^{\mathrm{Irr}(\mathcal{N})} \times \mathbb{R}_+^{2\mathrm{Rev}(\mathcal{N})}. \tag{1.12}$$

Define the *reduction map* $\Psi : \Gamma' \to \Gamma$ that identifies reversible reaction pairs in $\mathcal{N}'$ and defines the flux through the corresponding reversible reaction in $\mathcal{N}$ as the net flux trough this reaction. So for $\Phi' \in \Gamma'$ we have

$$\Psi(\Phi')_r := \begin{cases} \Phi'_r, & \text{if } r \in \mathrm{Irr}(\mathcal{N}), \\ \Phi'_{(r,+)} - \Phi'_{(r,-)}, & \text{if } r \in \mathrm{Rev}(\mathcal{N}). \end{cases} \tag{1.13}$$

The map Ψ is additive and positively homogeneous. It is surjective, but not injective generally. One can define a 'right-inverse', $\Psi^\dagger : \Gamma \to \Gamma'$ such that $\Psi \circ \Psi^\dagger = \mathrm{Id}_\Gamma$. We call this map *standard splitting* or *standard reconfiguration* of a flux distribution $\Phi \in \Gamma$ by assigning all *net* flux to the corresponding reaction in positive or negative direction:

$$\Psi^\dagger(\Phi)_r := \begin{cases} \Phi_i, & \text{if } r = i,\ i \in \mathrm{Irr}(\mathcal{N}), \\ \Phi_i, & \text{if } r = (i,+),\ i \in \mathrm{Rev}(\mathcal{N}),\ \text{and } \Phi_i \geq 0, \\ -\Phi_i, & \text{if } r = (i,-),\ i \in \mathrm{Rev}(\mathcal{N}),\ \text{and } \Phi_i \leq 0, \\ 0, & \text{otherwise.} \end{cases} \tag{1.14}$$

The stoichiometric matrix of the reconfigured network can now be expressed as

$$S'\Phi' = S\Psi(\Phi') \qquad \text{for } \Phi' \in \Gamma'. \tag{1.15}$$

Exercise 1.4.4 *Show that Ψ maps the steady state flux cone $(\Gamma')^*$ of $\mathcal{N}'$ onto the steady state flux cone Γ^* of $\mathcal{N}$: $\Gamma^* = \Psi\big((\Gamma')^*\big)$.*

Before we can state the main result that relates EFMs and ECs we need to introduce the concept of a futile cycle. A *futile cycle* in $\mathcal{N}'$ is a flux distribution $\Phi' \in (\Gamma')^*$ such that it involves a single reversible reaction pair and $\Psi(\Phi') = 0$, i.e. the fluxes in either reaction are the same. A futile cycle does not have any net production.

The main result that relates EFMs and ECs is now the following:

Theorem 1.4.5 (Relationship EFMs and ECs) *The following holds:*

(i) The standard reconfiguration of an EFM of $\mathcal{N}$ through $\Psi^\dagger$ is an EC of the reconfigured network $\mathcal{N}'$. Consequently, each EFM of $\mathcal{N}$ is the reduction of some EC of $\mathcal{N}'$ under Ψ.

(ii) An EC of $\mathcal{N}'$ is either the standard reconfiguration of an EFM of $\mathcal{N}$, or it is a futile cycle.

Thus, the EFMs of $\mathcal{N}$ are in one-to-one correspondence with the non-zero reductions of ECs of $\mathcal{N}'$. In particular, the set of EFMs is finite, generate Γ^* and is uniquely determined by $\mathcal{N}$.

Since $\Gamma^* = \Psi((\Gamma')^*)$ and $(\Gamma')^*$ is generated by the collection of its ECs (see Section 1.4.1), a consequence of Theorem 1.4.5 is, that Γ^* is generated by the collection of EFMs. Thus, any steady state flux mode of $\mathcal{N}$ can thus be viewed as non-negative linear combination of EFMs. This representation need not be unique.

E.g. [10] provides an algorithm for computing EFMs from the stoichiometric matrix that does not require the computation of ECs of the reconfigured network. In Figure 1.2 an example is provided of a network with reversible reactions, the associated reconfigured network and computed ECs and EFMs.

1.4.3 Extreme Pathways

Extreme Pathways were introduced by Palsson and coworkers [8, 5] independently from Schuster *et al.* at almost the same time to work on particular metabolic networks with reversible reactions. In there set-up there is a single interior compartment, all internal reactions therein are irreversible, while some exchange reactions are allowed to be reversible, with the particular condition, that each internal metabolite is involved in at most one reversible exchange reaction. We shall call this the 'EP-condition'

One may show that this condition implies that the steady state flux cone is pointed. Hence it is generated by its extreme rays (Theorem 1.4.1). The flux modes that correspond to these extreme rays are called *Extreme Pathways*

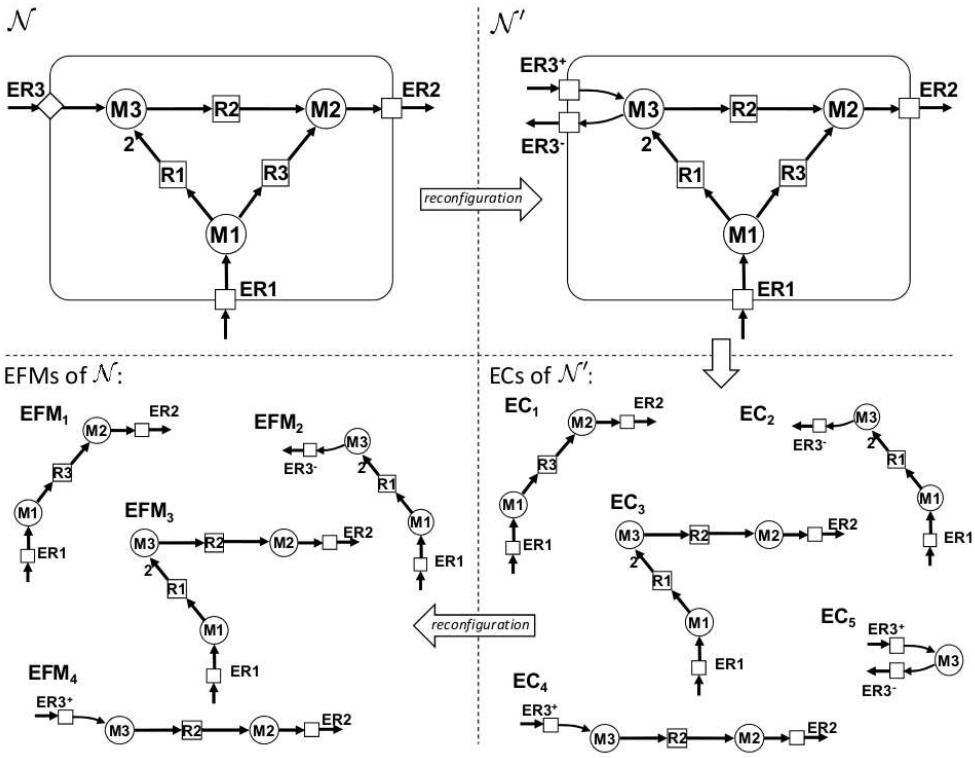

Figure 1.2: *An example of a reaction network $\mathcal{N}$ with reversible reactions and its reconfiguration $\mathcal{N}'$. The computed ECs of $\mathcal{N}'$ are presented as pathways.* EC_5 *is a futile cycle.* $\mathrm{EC}_1, \ldots, \mathrm{EC}_4$ *correspond to EFMs of network $\mathcal{N}$.*

(EPs). A network that satisfies the EP-condition will typically have more EFMS than EPs. The excess EFMs are necessarily non-negative linear combinations of the EPs, because the latter generate the steady state flux cone. Reaction network $\mathcal{N}$ in Figure 1.2 satisfies the EP-condition. EFM_1, EFM_2 and EFM_4 are EPs. EFM_3 is the combination of EFM_2 and EFM_4.

Allthough the EPs thus form a smaller collection of flux modes that still generates the steady state flux cone, their applicability is mainly limited by the EP-condition. Various real-life examples will fail the condition that an internal metabolites is involved in at most one reversible exchange reaction. Nevertheless, Palsson and coworkers have provided a wealth of applications of EP-analysis in a series of papers, see e.g. [1, 2, 6], among others.

1.5 Applications of Elementary Flux Modes

Once a sufficiently small biologically meaningful collection of flux modes is obtained that generates the steady state flux cone one can apply these for various purposes. The ideas for applications put forward by the Palsson school, using EPs for networks that satisfy the EP-condition, can equally well be performed with EFMs. As the collection of EFMs as generating set for the steady state flux cone exists for any chemical reaction network, we continue by considering

these as collection of generators. We mention a few applications:

One can determine *network statistics* taking the biological meaningful and *objective* metabolic pathways as defined by the EFMs as starting point. The pathways are obtained by mathematical means from the stoichiometry of the network rather than by subjective inspection. Using the EFMs one may consider e.g. the distribution of pathway lenghts, pathway overlap, correlated reaction subsets, ...

It allows *comparison among different organisms*. Network-based statistics as above for different organisms can be used to compare their metabolic networks. Observed differences in network statistics may require further biological research for explanation and understanding.

One can perform *in silico bioengineering experiments* to see what would be the effect of a modification of the metabolic network before this is achieved in a living organism (bacterium, yeast). An example is provided in [1] (using EPs).

Biologically, the number of representations of a given flux mode that involve different sets of EFMs provides *quantification of the versatility of the organism* in organising its metabolism.

Analyse the possible ways of *functioning of the metabolic network*: one can count the number of EFMs that are able to produce a given product from a given substrate. Each EFM has an *input-output signature*: a set of substrates that is taken up from the environment that results in a set of products. Inspection of the possible input-output signatures gives insight in the capabilities of the network.

One can show that optimal yield of a product takes effect in an EFM. So one can identify optimal environmental conditions under which optimal production is possible.

Phenotypic Phase Plane Analysis allows to analyse what combination of EFMs is needed to have maximal production rate of a compound *for a given pair of influx rates* for two substrates from the environment, and changes in these combinations. This provides insight into changes in gene regulation that are needed for the organism to optimally perform (i.e. with regard to production of the chosen product) when its environment changes.

1.6 Flux Balance Analysis

1.7 Reaction Network Theory

Reaction Network Theory is a branch of mathematics that is concerned with the question to what extent the possible long-term dynamics of a reaction network is determined by *network topology*, i.e. the way in which the reaction are

coupled together, independent of parameter values. A notable example thereof is Feinberg's '*Deficiency-One Theorem*' [4]. If all reactions fluxes are given by Mass Action Kinetics, then a particular topology of the connections (the 'Deficiency-one') implies that there can exist a single steady state only.

Application to metabolic networks seems of such theory seems limited, because of the Mass Action assumption. Most reactions in a metabolic network are enzyme-catalyzed, so reaction rates do not satisfy this assumption when the enzyme and enzyme-substrate complex(es) are not included in the model. That is, when effective rate expressions for the reactions are used, obtained by suitable time scale separation arguments (See Chapter 2).

Bibliography

[1] Carlson, R., Fell, D. and F. Srienc (2002). Metabolic pathway analysis of a recombinant yeast for rational strain development, *Biotech. Bioeng.* **79**(2), 121–134.

[2] Covert M.W. and B.Ø. Palsson (2003). Constraints-based models: regulation of gene expression reduces the steady-state solution space, *J. Theor. Biol.* **221**, 309–325.

[3] Clarke, B.L. (1988). Stoichiometric network analysis, *Cell Biophysics* **12**(1), 237–253.

[4] Feinberg, M. (1995), The existence and uniqueness of steady states for a class of chemical reaction networks, *Arch. Rational Mech. Anal.* **132**, 311–370.

[5] Palsson, B.Ø. (2006). *Systems Biology: Properties of reconstructed networks*, Cambridge: Cambridge University Press.

[6] Papin, J.A., Price N.D. and B.Ø. Palsson (2002). Extreme patway lengths and reaction participation in genome-scale metabolic networks, *Genome Res.* **12**, 1889–1900.

[7] Rockafellar, R.T. (1970). *Convex Analysis*, Princeton: Princeton University Press.

[8] Schilling, Ch.L., Letscher, D. and B.ØPalsson (2000). Theory for the systemic definition of metabolic pathway and their use in interpreting metabolic function from a pathway-oriented perspective, *J. Theor. Biol.* **203**, 229–248.

[9] Schuster, S., Fell, D.A. and Th. Dandekar (2000). A general definition of metabolic pathways useful for systematic organization and analysis of complex metabolic networks, *Nature Biotech.* **18**, 326–332.

[10] Schuster, S., Hilgetag, C., Woods, J.H. and D.A. Fell (2002). Reaction routes in biochemical reaction systems: algebraic properties, validated calcuation procedures and example from nucleotide metabolism, *J. Math. Biol.* **45**(2), 153–182.