

Contents lists available at ScienceDirect

Discrete Applied Mathematics

journal homepage: www.elsevier.com/locate/dam



A new constraint-based description of the steady-state flux cone of metabolic networks

Abdelhalim Larhlimi, Alexander Bockmayr*

DFG-Research Center Matheon, FB Mathematik und Informatik, Freie Universität Berlin, Arnimallee 3, D-14195 Berlin, Germany

ARTICLE INFO

Article history:
Received 8 January 2007
Received in revised form 24 September 2007
Accepted 11 June 2008
Available online 9 September 2008

Keywords: Metabolic network Metabolic pathway analysis Polyhedral cone Steady-state flux cone

ABSTRACT

Metabolic pathway analysis is an important area in computational biology, which has received considerable attention in the recent past. The set of all possible flux distributions over a metabolic network at steady state defines a polyhedral cone, the steady-state flux cone. Two major approaches exist to characterize this cone: elementary flux modes and extreme pathways. Both use an inner description of the flux cone, which is based on sets of generating vectors. The number of these generators may be very large even for small networks. This limits the practical applicability of these methods.

We present a new constraint-based approach to metabolic pathway analysis which uses an outer description of the flux cone, based on sets of non-negativity constraints. These can be identified with irreversible reactions and thus have a direct biochemical interpretation. The resulting description of the flux cone is minimal and unique. Furthermore, it satisfies a simplicity condition similar to the one that holds for elementary flux modes.

© 2008 Elsevier B.V. All rights reserved.

1. Introduction

Metabolic pathway analysis is an important area in computational biology, which has received considerable attention in the recent past [10,12,13,18]. The constraints that have to hold in a metabolic network at steady state, which include stoichiometry and thermodynamic irreversibility, have led to the definition of the steady-state flux cone, which contains all the possible flux distributions over the network. Two important concepts have been proposed to describe the flux cone in a mathematically and biologically meaningful way: elementary flux modes and extreme pathways, see e.g. [6,9,16,19,20, 22]. These two approaches are closely related. Both use an inner description of the flux cone based on generating vectors. Furthermore, the extreme pathways are a subset of the elementary flux modes [9]. However, from a mathematical point of view, not all of the elementary flux modes or extreme pathways are needed to describe the cone. This observation is important because the computation of these vectors in a high-dimensional space often leads to a combinatorial explosion [8,13]. Even for small networks, the number of elementary flux modes or extreme pathways may be very large. This makes a complete analysis of the whole network impossible and limits the practical applicability of these methods.

In this paper, we propose a new constraint-based approach to metabolic pathway analysis, which uses an outer description of the steady-state flux cone, based on sets of non-negativity constraints. These can be identified with irreversible reactions and therefore have a direct biochemical interpretation. Our method is thus different from existing approaches, such as elementary flux modes or extreme pathways, which use an inner description. We characterize a metabolic network by two new concepts: *minimal metabolic behaviors* (MMBs) and the *reversible metabolic space* (RMS). Like elementary flux modes or extreme pathways, these are uniquely determined by the network. The set of all MMBs together with the RMS yields a complete description of the flux cone, which is minimal, unique, and satisfies a simplicity condition similar to the

^{*} Corresponding author. Fax: +49 30 838 75869.

E-mail addresses: larhlimi@mi.fu-berlin.de (A. Larhlimi), bockmayr@mi.fu-berlin.de (A. Bockmayr).

one that holds for elementary flux modes. Moreover, our approach leads to a new classification of reactions (irreversible, pseudo-irreversible, fully reversible), which may be used for a refined analysis of the network.

The organization of this paper is as follows. We start in Section 2 with some basic facts about polyhedral cones. In Section 3, we recall metabolic pathway analysis and give a first presentation of our approach. In Section 4, we formally introduce minimal metabolic behaviors and the reversible metabolic space. Section 5 contains a classification of reactions according to their reversibility type. It also describes how MMBs can be used to decompose a given metabolic network into functional subnetworks. Finally, in Section 6, we present some computational results in order to compare MMBs and the RMS with extreme pathways and elementary modes.

2. Polyhedral cones

We start with some basic facts about polyhedral cones (see e.g. [17]).

A non-empty subset $C \subseteq \mathbb{R}^n$ is called a *(convex) cone* if $\lambda x + \mu y \in C$, whenever $x, y \in C$ and $\lambda, \mu \geq 0$. A cone C is *polyhedral*, if $C = \{x \in \mathbb{R}^n \mid Ax \geq 0\}$, for some real matrix $A \in \mathbb{R}^{m \times n}$. If this is the case, lin.space(C) = $\{x \in \mathbb{R}^n \mid Ax = 0\}$ is called the *lineality space* of C.

A cone *C* is *finitely generated* if there exist $g^1, \ldots, g^s \in \mathbb{R}^n$ such that $C = \text{cone}\{g^1, \ldots, g^s\} \stackrel{\text{def}}{=} \{\lambda_1 g^1 + \cdots + \lambda_s g^s \mid \lambda_1, \ldots, \lambda_s \geq 0\}$. A fundamental theorem of Farkas–Minkowski–Weyl (see e.g. [17], p. 87) asserts that a convex cone is polyhedral if and only if it is finitely generated. For the rest of this paper, we will consider only polyhedral cones.

An inequality $a^Tx \ge 0$, $a \in \mathbb{R}^n \setminus \{0\}$, is valid for C if $C \subseteq \{x \in \mathbb{R}^n \mid a^Tx \ge 0\}$. The set $F = C \cap \{x \in \mathbb{R}^n \mid a^Tx = 0\}$ is then called a *face* of C. The *dimension* of F is defined as the dimension of the linear subspace generated by F.

Any non-zero element $r \in C$ is called a *ray* of C. Two rays r and r' are equivalent, written $r \cong r'$, if $r = \lambda r'$, for some $\lambda > 0$. A ray r is *extreme* if there do not exist rays r', $r'' \in C$, $r' \ncong r''$, such that r = r' + r''.

Pointed cones. A polyhedral cone C is called *pointed* if $\lim_{C \to C} S(C) = \{0\}$. Any pointed cone C has a canonical representation

$$C = \operatorname{cone}\{r^1, \dots, r^s\},\tag{1}$$

where r^1, \ldots, r^s are the (distinct) extreme rays of C. This representation, which is used in the extreme pathway approach, is minimal and unique up to multiplication with positive scalars.

Non-pointed cones. If C is not pointed, there is no longer such a unique minimal representation. Let t be the dimension of the lineality space of C. Instead of the extreme rays, we consider now the minimal proper faces G^1, \ldots, G^s of C, which are the

faces of *C* of dimension t + 1. Each G^i can be represented by a row vector a_i^T and a submatrix A_i' of *A*, with rank $\binom{A_i'}{a_i^T} = n - t$, such that [17]

$$G^{i} = \{x \in C \mid a_{i}^{T} x \ge 0, A_{i}' x = 0\}, \tag{2}$$

and

lin.space(
$$C$$
) = { $x \in C \mid a_i^T x = 0, A_i' x = 0$ }.

If we select for each $i=1,\ldots,s$ a vector $g^i\in G^i\setminus \mathrm{lin.space}(C)$, and vectors $b^0,\ldots,b^t\in \mathrm{lin.space}(C)$ such that $\mathrm{lin.space}(C)=\mathrm{cone}\{b^0,\ldots,b^t\}$, we get

$$C = \operatorname{cone}\{g^1, \dots, g^s, b^0, \dots, b^t\}. \tag{3}$$

For each minimal proper face G^i , i = 1, ..., s, we get

$$G^{i} = \operatorname{cone}\{g^{i}\} + \operatorname{lin.space}(C) = \{\lambda g^{i} + w \mid \lambda \ge 0, w \in \operatorname{lin.space}(C)\}. \tag{4}$$

For additional information, we refer to [17], p. 105–106.

(3) generalizes (1), but this representation is no longer unique. We may choose an arbitrary base of lin.space(C), and arbitrary rays g^i in $G^i \setminus \text{lin.space}(C)$. However, it follows from (2) that G^i can also be characterized using constraints $a_i^T x \ge 0$, where a_i^T is a row vector from the matrix A that defines the cone. This observation leads us to a new way for describing and analyzing the flux cone associated with a metabolic network, which is developed in this paper.

3. Metabolic pathway analysis

In the context of metabolic pathway analysis, metabolic systems are assumed to operate at steady state so that the rate of production and the rate of consumption of each internal metabolite must be equal. In addition, the flux through each irreversible reaction must be non-negative. Mathematically, the stoichiometric and thermodynamic constraints that have to hold in a metabolic network can be expressed as follows [13]:

$$Sv = 0, \quad v_i \ge 0, \text{ for all } i \in Irr,$$
 (5)

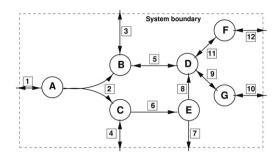


Fig. 1. Network example ILLUSNET.

where *S* is the $m \times n$ stoichiometric matrix of the network, with m internal metabolites (rows) and n reactions (columns), and $v \in \mathbb{R}^n$ is the flux vector. Irr $\subseteq \{1, \ldots, n\}$ denotes the set of *irreversible* reactions in the network, and Rev $= \{1, \ldots, n\} \setminus \text{Irr}$ is the set of *reversible* reactions.

When modeling a metabolic system, we distinguish between *external* and *internal* metabolites [7]. External metabolites are the sources and sinks of the network. For the internal metabolites, we assume that there is no accumulation or depletion at steady state. Thus, the rate of production must equal the rate of consumption. In general, the classification of a metabolite as external or internal depends on the purpose of the model. It should be noted that this classification has also an impact on the algorithmic complexity of analyzing the network [4].

The set of all solutions of the constraint system (5), which corresponds to the set of all possible flux distributions over the network at steady state, defines a polyhedral cone,

$$C = \{ v \in \mathbb{R}^n \mid Sv = 0, \ v_i \ge 0, \text{ for all } i \in Irr \}$$
 (6)

which is called the *flux cone* [2,3]. Already in [2], we can find the distinction between inner and outer descriptions of this cone, which are called there internal and external representations. The external representation gives a test for determining whether a given flux vector belongs to the cone, while the internal representation allows one to construct flux vectors from a set of generators.

Example 1. For illustration throughout this paper, we consider the hypothetical network ILLUSNET depicted in Fig. 1. It consists of seven metabolites (A, \ldots, G) , and twelve reactions $(1, \ldots, 12)$. The steady-state flux cone is defined by $C = \{v \in \mathbb{R}^{12} \mid Sv = 0, v_i \geq 0, \text{ for all } i \in Irr\}$, with the set of irreversible reactions $Irr = \{2, 6, 7, 8\}$, and the stoichiometric matrix

If all reactions in a metabolic network are irreversible (and also in some other cases), the flux cone is pointed, and has a minimal and unique set of generators, which correspond to the extreme rays. In general, however, the flux cone is not pointed, and there is no such a canonical representation.

In the elementary mode approach, a unique set of generators is obtained by imposing a simplicity condition [20]. An elementary flux mode corresponds to a flux vector $v \in C$ such that there is no $v' \in C \setminus \{0\}$ with $S^c(v') \stackrel{\text{def}}{=} \{i \mid v_i' \neq 0\} \subsetneq S^c(v) \stackrel{\text{def}}{=} \{i \mid v_i \neq 0\}$. This yields a complete set of generators of the cone, which however is not minimal.

In the extreme pathway approach, the network is reconfigured and all internal reversible reactions are split into a forward and a backward reaction, where both are constrained to be non-negative [16]. This implies that the flux cone becomes pointed and thus has a minimal and unique (up to multiplication with positive scalars) set of generators. These correspond to the extreme rays of the reconfigured cone. However, this reconfiguration has undesirable consequences. On the one hand, the number of variables and constraints increases, and the constraint system becomes more difficult to solve. On the other hand, a significant number of pathways in the reconfigured cone are extreme for the only reason that the internal reversible reactions have been decomposed into forward and backward reactions. In the initial cone, these extreme pathways are conically dependent. In addition, the number of extreme pathways (and even more the number of elementary modes) is often very large, even for small networks. This limits the practical applicability of the approach. Computational results illustrating these observations will be given in Section 6.

To overcome these problems, we propose a new description of the flux cone. Mathematically speaking, we switch from an inner description of the cone, based on generating vectors, to an outer description, based on constraints. In our case,

these constraints are inequalities of the form $v_i \geq 0$, $i \in Irr$, which can be identified with irreversible reactions, and thus have an immediate biochemical interpretation. Working directly on a possibly non-pointed flux cone C, we associate with each minimal proper face a characteristic set of irreversible reactions, which corresponds to a minimal metabolic behavior (MMB). In addition, we include in our description the lineality space lin.space(C), which will be called the reversible metabolic space (RMS). For a non-pointed cone, this space is different from $\{0\}$. It may contain useful biological information, which is no longer explicit if we replace a reversible reaction with two irreversible ones. The set of all MMBs together with the RMS yields a complete description of the flux cone, which is minimal, unique, and satisfies a simplicity condition similar to the one that holds for elementary flux modes. In addition, by distinguishing three types of reactions (irreversible, pseudo-irreversible, fully reversible), this approach provides a different view of the network, which may be used in a refined analysis.

4. Minimal metabolic behaviors

In this section, we formally introduce minimal metabolic behaviors and the reversible metabolic space. If the metabolic network does not contain any irreversible reaction, the steady-state flux cone becomes a linear subspace of \mathbb{R}^n , which can be analyzed by standard methods from linear algebra. Therefore, we assume for the rest of this paper that the metabolic network contains at least one irreversible reaction.

4.1. Characterizing minimal proper faces

We start by characterizing the minimal proper faces of the flux cone through irreversible reactions of the network.

Definition 2. Let G be a minimal proper face of the flux cone C and let $j \in Irr$ be an irreversible reaction. We say that G is *characterized by j* if there exists $I_j \subset Irr$ such that $G = \{v \in C \mid v_j \geq 0, v_i = 0, \text{ for all } i \in I_j\}$, and lin.space(C) = $\{v \in C \mid v_i = 0, v_i = 0, \text{ for all } i \in I_j\}$.

It follows from Eq. (2) that each minimal proper face *G* of *C* is characterized by at least one irreversible reaction. However, this reaction need not be unique. In general, there will be several irreversible reactions satisfying the conditions of Definition 2. The following proposition provides a simple criterion to identify the irreversible reactions that characterize a given minimal proper face.

Proposition 3. Let G be a minimal proper face of C and let $j \in Irr$ be an irreversible reaction. Then the following are equivalent:

- (1) G is characterized by i.
- (2) $v_j > 0$, for some $v \in G \setminus \text{lin.space}(C)$.
- (3) $v_i > 0$, for all $v \in G \setminus \text{lin.space}(C)$.

Proof. (1) \Rightarrow (2): Since dim(G) = 1 + dim(lin.space(C)), we have $G \setminus \text{lin.space}(C) \neq \emptyset$. So there exists $v \in G \setminus \text{lin.space}(C)$, with $v_i > 0$.

- $(2)\Rightarrow (3)$: Suppose $g\in G\setminus \text{lin.space}(C)$ with $g_j>0$. By Eq. (4), for any $v\in G\setminus \text{lin.space}(C)$ there exist $\lambda>0$ and $w\in \text{lin.space}(C)$ such that $v=\lambda\cdot g+w$. It follows that $v_j=\lambda\cdot g_j>0$.
- (3) \Rightarrow (1): It follows from Eq. (2) that G is characterized by at least one irreversible reaction. So there exist $k \in Irr$ and $I_k \subset Irr$ with $G = \{v \in C \mid v_k \geq 0, \ v_i = 0, \ \text{for all } i \in I_k\}$, and $\lim_{s \to a} (C) = \{v \in C \mid v_k = 0, \ v_i = 0, \ \text{for all } i \in I_k\}$. To prove (1), we set $I_j = I_k$ and claim that the same equations hold for k replaced with j. Consider $v \in C$ with $v_i = 0$, for all $i \in I_j = I_k$. Since $v \in C$, we have $v_j \geq 0$ and $v_k \geq 0$. If $v_k > 0$, then $v \in G \setminus \lim_{s \to a} (C)$ and by (3) $v_j > 0$. If $v_k = 0$, then $v \in I_k$ and $v \in$

The proposition motivates our next definition.

Definition 4. Given a minimal proper face G of the flux cone C, the set

```
D = \{j \in Irr \mid v_i > 0, \text{ for some } v \in G\}
```

of all irreversible reactions characterizing *G* is called the *characteristic set* of *G*.

Each minimal proper face G of the flux cone is thus characterized by a set D of irreversible reactions. As the next theorem shows, all flux vectors $v \in G \setminus \text{lin.space}(C)$ have the following common property: the flux through all irreversible reactions belonging to D is positive, i.e., $v_j > 0$, for $j \in D$, while the flux through all the other irreversible reactions is zero, i.e., $v_j = 0$, for $j \in \text{Irr} \setminus D$.

Since $\lim_{C \to \infty} (C) = \{v \in C \mid v_i = 0 \text{ for all } i \in \operatorname{Irr} \}$, flux vectors in $\lim_{C \to \infty} (C)$ involve only reversible reactions. Hence, a vector $v \in \lim_{C \to \infty} (C)$ is itself reversible, i.e., $-v \in \lim_{C \to \infty} (C)$. Note that information about reversible pathways is lost if the network is reconfigured in order to obtain a pointed cone.

Theorem 5. Let G be a minimal proper face of the flux cone C and D its characteristic set. Then

$$G = \{v \in C \mid v_i > 0, \text{ for all } j \in D, v_i = 0, \text{ for all } i \in Irr \setminus D\} \cup lin.space(C).$$

Proof. Suppose $j \in D$. Then $G = \{v \in C \mid v_j \geq 0, \ v_i = 0, \ \text{for} \ i \in I_j\}$ and $\text{lin.space}(C) = \{v \in C \mid v_j = 0, \ v_i = 0, \ \text{for} \ i \in I_j\}$, for some $I_j \subset \text{Irr}$. From Proposition 3, we see that $I_j \subset \text{Irr} \setminus D$. It follows that $\{v \in C \mid v_j > 0, \ \text{for} \ \text{all} \ j \in D, \ v_i = 0, \ \text{for} \ \text{all} \ i \in \text{Irr} \setminus D\}$. Unin.space(C) $\subseteq G$. To show the reverse inclusion, suppose $v \in G \setminus \text{lin.space}(C)$. Then, following Proposition 3, $v_j > 0$, for all $j \in D$. Suppose $v_i > 0$, for some $i \in \text{Irr} \setminus D$. From Definition 4, we would get $i \in D$, which is a contradiction. \square

Note that the set D is uniquely determined by G. If G^1, \ldots, G^s are the minimal proper faces of the flux cone C, the corresponding characteristic sets D^1, \ldots, D^s together with the lineality space lin.space(C) completely describe C.

The next result shows that inside a minimal proper face G, the fluxes through the irreversible reactions in D are proportional to each other.

Corollary 6. Let D be the characteristic set of the minimal proper face G. Then for all $j, k \in D$, there exists $\alpha > 0$ such that $v_k = \alpha \cdot v_i$, for all $v \in G$. In particular, $v_i = 0$ implies $v_k = 0$, and $v_i > 0$ implies $v_k > 0$, for all $v \in G$.

Proof. Consider $g \in G \setminus \text{lin.space}(C)$. Since $j, k \in D$, Proposition 3 implies $g_j > 0$ and $g_k > 0$. By Eq. (4), for all $v \in G \setminus \text{lin.space}(C)$, there exist $\lambda > 0$ and $w \in \text{lin.space}(C)$ such that $v = \lambda \cdot g + w$. It follows that $v_j = \lambda \cdot g_j > 0$, $v_k = \lambda \cdot g_k > 0$, and therefore $v_j/v_k = g_j/g_k \stackrel{\text{def}}{=} \alpha > 0$, independently from v. This shows that $v_j = \alpha \cdot v_k > 0$, for all $v \in G \setminus \text{lin.space}(C)$. For all $v \in G \setminus \text{lin.space}(C)$, we have $v_j = v_k = 0$. It follows for all $v \in G \setminus \text{lin.space}(C)$.

4.2. Minimal metabolic behaviors and the reversible metabolic space

We are now ready to define the key notions of this paper.

Definition 7. A metabolic behavior is a set of irreversible reactions $D \subseteq \operatorname{Irr}, D \neq \emptyset$, such that there exists a flux vector $v \in C$ with

$$D = \{i \in \operatorname{Irr} \mid v_i \neq 0\}. \tag{7}$$

A metabolic behavior *D* is *minimal*, if there is no metabolic behavior $D' \subseteq D$ strictly contained in *D*. The set

$$\{v \in C \mid v_i = 0, \text{ for all } i \in Irr\}$$
 (8)

is called the reversible metabolic space.

Remember that elementary flux modes correspond to flux vectors $v \in C$ involving a minimum set of reactions, i.e., the set $S^c(v) = \{i \in \text{Rev} \cup \text{Irr} \mid v_i \neq 0\}$ is minimal [20]. Similarly, a minimal metabolic behavior corresponds to a minimal set of *irreversible* reactions involved in a flux vector $v \in C \setminus \{0\}$, i.e., the set $D = \{i \in \text{Irr} \mid v_i \neq 0\}$ is minimal.

Proposition 8. A set $D \subseteq Irr$, $D \neq \emptyset$, is a minimal metabolic behavior if and only if the following two conditions hold:

- (1) There exists $v \in C$ with $v_i > 0$, for all $i \in D$, and $v_i = 0$, for all $i \in Irr \setminus D$.
- (2) For any $v \in C$ with $v_i = 0$ for all $i \in Irr \setminus D$, if $v_i = 0$ for some $j \in D$, then $v_i = 0$ for all $j \in D$.

Proof. " \Rightarrow ": Suppose $D \subseteq Irr$ is a minimal metabolic behavior. Since D is a metabolic behavior, there exists $v \in C$ such that $D = \{i \in Irr \mid v_i \neq 0\}$. Since $v \in C$, we get $v_i > 0$, for all $i \in D$. By definition of D, $v_i = 0$, for all $i \in Irr \setminus D$. This shows (1). Now let $v \in C$ with $v_i = 0$ for all $i \in Irr \setminus D$. Suppose $v_j = 0$ for some $j \in D$ and $v_k \neq 0$, for some $k \in D$, $k \neq j$. Then $D' = \{i \in Irr \mid v_i \neq 0\}$ is a metabolic behavior strictly contained in D, contradicting the minimality of D.

"\(\infty\)": Consider $\emptyset \neq D \subseteq \operatorname{Irr}$ such that (1) and (2) hold. By (1), there exists $v \in C$ with $v_i > 0$, for all $i \in D$, and $v_i = 0$, for all $i \in \operatorname{Irr} \setminus D$. Then $D = \{i \in \operatorname{Irr} \mid v_i \neq 0\}$ and so D is a metabolic behavior. Suppose D is not minimal. Then there exists a metabolic behavior $\emptyset \neq D' \subsetneq D$ strictly contained in D. Let $D' = \{i \in \operatorname{Irr} \mid v_i' \neq 0\}$, for a suitable $v' \in C$. Then $v_i' = 0$, for $i \in \operatorname{Irr} \setminus D' \supseteq \operatorname{Irr} \setminus D$. Since $D' \subsetneq D$, there exists $i \in D$ with $v_j' = 0$. From (2) we get $v_j' = 0$, for all $j \in D$. This is a contradiction, since $D' \neq \emptyset$ implies that there exists $i \in D' \subseteq D$ with $v_i' \neq 0$. \square

The following theorem shows that the MMBs are in a 1-1 correspondence with the minimal proper faces of the flux cone. Indeed, each minimal metabolic behavior is identical to the characteristic set of a minimal proper face.

Theorem 9. Let $D \subseteq Irr$ be a set of irreversible reactions. Then, the following are equivalent:

- D is a minimal metabolic behavior.
- There exists a minimal proper face G whose characteristic set is D.

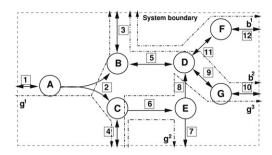


Fig. 2. Representative pathways in ILLUSNET.

Proof. " \Rightarrow ": Suppose D is a minimal metabolic behavior and let $G = \{v \in C \mid v_j \geq 0, \text{ for all } j \in D, v_i = 0, \text{ for all } i \in Irr \setminus D\}$. Since $G = \{v \in C \mid v_i = 0, \text{ for all } i \in Irr \setminus D\}$, G is a face of C (cf. [17], p. 101). Let $G' \subseteq G$ be a minimal proper face of C and D' its characteristic set. Since $G' \subseteq G$, we get $D' \subseteq D$. Suppose there exists $k \in D \setminus D'$. Then $v_k = 0$ for all $v \in G'$ and $G' \subseteq G \cap \{v \in \mathbb{R}^n \mid v_k = 0\}$. Since D is minimal, $v \in G$ and $v_k = 0$ implies $v_j = 0$, for all $v_k \in D'$ and therefore $v_k \in C' \cap \{v \in \mathbb{R}^n \mid v_k = 0\}$ = lin.space(C). It follows that $C' \subseteq C' \cap \{v \in \mathbb{R}^n \mid v_k = 0\}$ = lin.space(C). Applying Theorem 5 to C', we get $C \subseteq C' \cap C' \cap C'$ and $C \subseteq C' \cap C' \cap C' \cap C'$ and $C \subseteq C' \cap C' \cap C' \cap C'$ and $C \subseteq C' \cap C' \cap C' \cap C'$ and $C \subseteq C'$ and $C \subseteq C' \cap C'$ and $C \subseteq C'$ and

" \Leftarrow ": Let G be a minimal proper face with characteristic set D. We use Proposition 8 to show that D is a minimal metabolic behavior. If $v \in G \setminus \text{lin.space}(C)$, then by Theorem 5, we get $v_i > 0$ for all $i \in D$ and $v_i = 0$ for all $i \in \text{Irr} \setminus D$, i.e., v satisfies condition (1) of Proposition 8. To check condition (2), let $v \in C$ such that $v_i = 0$ for all $i \in \text{Irr} \setminus D$. From Theorem 5, we get $v \in G$. If $v_i = 0$ for some $j \in D$, Corollary 6 yields $v_k = 0$ for all $k \in D$. \Box

Example 10. In the metabolic network from Fig. 2, the MMBs and the corresponding minimal proper faces are as follows:

$$\begin{split} D^1 &= \{2\}, \qquad D^2 &= \{6,7\}, \qquad D^3 &= \{6,8\}, \\ G^k &= \{v \in C \mid v_j \geq 0, j \in D^k, \ v_i = 0, i \in \operatorname{Irr} \setminus D^k\}, \quad k = 1,2,3. \end{split}$$

Note that the irreversible reaction 6 is participating in the definition of two minimal proper faces, G^2 and G^3 . Fig. 2 shows three pathways

```
\begin{split} g^1 &= (1, 1, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0), \\ g^2 &= (0, 0, 0, -1, 0, 1, 1, 0, 0, 0, 0, 0), \\ g^3 &= (0, 0, 0, -1, 0, 1, 0, 1, 1, 1, 0, 0), \end{split}
```

representing the minimal proper faces G^1 , G^2 , and G^3 , respectively. The reversible metabolic space lin.space(C) = { $v \in C \mid v_i = 0, i \in Irr$ } has dimension 2. As a vector space, it can be generated by the pathways

```
b^1 = (0, 0, -1, 0, 1, 0, 0, 0, 0, 0, 1, 1),

b^2 = (0, 0, -1, 0, 1, 0, 0, 0, 1, 1, 0, 0).
```

An arbitrary flux vector $v \in C$ can be written as linear combination $v = \sum_{k=1}^{3} \lambda_k g^k + \sum_{l=1}^{2} \mu_l b^l$, with $\lambda_k \ge 0$ and $\mu_l \in \mathbb{R}$. In this example, the number of elementary flux modes is 18, while the number of extreme pathways (after reconfiguration) is 14.

The irreversible reactions defining an MMB D cannot necessarily operate on their own. However, for each minimal metabolic behavior D, there exists at least one elementary flux mode involving exactly the irreversible reactions from D. For $v \in C$, let $D(v) = \{i \in Irr \mid v_i \neq 0\}$.

Proposition 11. Let D be a minimal metabolic behavior. Then there exists an elementary flux mode f such that D(f) = D.

Proof. Let D be a minimal metabolic behavior. According to Theorem 9, there exists a minimal proper face G whose characteristic set is D. Suppose $g \in G \setminus \text{lin.space}(C)$. Since the elementary flux modes generate C [20], $g = \sum_k \lambda_k f^k$ is a linear combination of elementary modes f^k , for some $\lambda_k \geq 0$. Since $g \neq 0$, there exists at least one elementary mode f^l such that $\lambda_l > 0$. For each $i \in D(f^l)$, we have $g_i = \sum_k \lambda_k f_i^k \geq \lambda_l f_i^l > 0$ and so $i \in D(g)$. This shows $D(f^l) \subseteq D(g)$. Since $g \in G$, it follows from Theorem 5 that D(g) = D. Finally, since D is a minimal metabolic behavior and $D(f^l) \subseteq D(g) = D$, we get $D(f^l) = D$ and the result follows. \square

In general, there can be more than one elementary $\operatorname{mode} f$ with D(f) = D. If this is the case, there are different elementary $\operatorname{modes} (\operatorname{possibly} \operatorname{many})$ that all belong to the same minimal proper face. In addition, there may exist elementary $\operatorname{modes} (\operatorname{possibly} \operatorname{many})$ that all belong to the same minimal proper face. In addition, there may exist elementary $\operatorname{modes} (\operatorname{possibly} \operatorname{many})$ in the interior of the flux cone C. We refer to Section 6 for computational results illustrating these remarks.

5. Pseudo-irreversible and fully reversible reactions

In this section, we classify reactions according to their reversibility type. We obtain a unique sign pattern for each MMB. These sign patterns can be used to decompose the network into minimal functional subnetworks (see Section 5.2).

5.1. Classification of reactions

We start by distinguishing two classes of reversible reactions.

Definition 12. Given the flux cone *C*, the set

$$Rev^0 = \{i \in Rev \mid v_i = 0, \text{ for all } v \in lin.space(C)\}\$$

is called the set of pseudo-irreversible reactions. Reactions in Frev = Rev \ Rev⁰ are called fully reversible.

Example 13. In the ILLUSNET network from Fig. 2, there is no reversible flux distribution involving reaction 4. We have $v_4 = 0$, for all $v \in \text{lin.space}(C)$. Therefore, reaction 4 is pseudo-irreversible. On the other hand, reaction 3 is involved in the reversible flux distribution b^1 which belongs to the lineality space lin.space(C). Thus, reaction 3 is fully reversible.

The next proposition shows that pseudo-irreversible reactions become irreversible inside minimal proper faces. Within each minimal proper face G, any pseudo-irreversible reaction with non-zero flux will take a unique direction, which is imposed by the MMB D associated with G. By taking the conical hull of the corresponding faces, we can identify a subspace of the cone in which the given pseudo-irreversible reaction takes only one direction.

Proposition 14. Let G be a minimal proper face of C and let $i \in \text{Rev}^0$ be a pseudo-irreversible reaction. Then exactly one of the following three conditions holds:

- (1) $v_i > 0$, for all $v \in G \setminus \text{lin.space}(C)$.
- (2) $v_i = 0$, for all $v \in G \setminus \text{lin.space}(C)$. (3) $v_i < 0$, for all $v \in G \setminus \text{lin.space}(C)$.

Proof. Suppose $g \in G \setminus \text{lin.space}(C)$. For any $v \in G \setminus \text{lin.space}(C)$ there exist $\lambda > 0$ and $w \in \text{lin.space}(C)$ such that $v = \lambda \cdot g + w$. Since $i \in \text{Rev}^0$, it follows that $sign(v_i) = sign(\lambda \cdot g_i) = sign(g_i)$, independently from v.

Example 15. In the ILLUSNET network, we have $Rev^0 = \{1, 4\}$, and $Frev = \{3, 5, 9, 10, 11, 12\}$. In the context of the MMB D^1 , the pseudo-irreversible reaction 4 becomes positive, i.e., $v_4 > 0$, for all $v \in G^1 \setminus \text{lin.space}(C)$, while it becomes negative in the context of D^2 and D^3 . The flux through the pseudo-irreversible reaction 1 is positive in D^1 , and zero in D^2 and D^3 .

The next result shows that inside a minimal proper face G, the fluxes through the pseudo-irreversible reactions involved in flux vectors $v \in G \setminus \text{lin.space}(C)$ are proportional to each other.

Corollary 16. Let G be a minimal proper face and let $j, k \in \text{Rev}^0$ be two pseudo-irreversible reactions. If there exist $g, g' \in \text{Rev}^0$ $G \setminus \text{lin.space}(C)$ such that $g_i \neq 0$ and $g'_k \neq 0$, then there exists $\alpha \neq 0$ with $v_k = \alpha \cdot v_i$, for all $v \in G$.

Proof. Suppose there exist $g, g' \in G \setminus \text{lin.space}(C)$ such that $g_j \neq 0$ and $g_k' \neq 0$. According to Proposition 14, we have $g_k \neq 0$. By Eq. (4), for all $v \in G \setminus \text{lin.space}(C)$, there exist $\lambda > 0$ and $w \in \text{lin.space}(C)$ such that $v = \lambda \cdot g + w$. Since $j,k \in \text{Rev}^0$, we get $v_j = \lambda \cdot g_j \neq 0$, $v_k = \lambda \cdot g_k \neq 0$, and therefore $v_k/v_j = g_k/g_j \stackrel{\text{def}}{=} \alpha \neq 0$, independently from v. This shows that $v_k = \alpha \cdot v_j \neq 0$, for all $v \in G \setminus \text{lin.space}(C)$. For all $v \in G \setminus \text{lin.space}(C)$, we have $v_j = v_k = 0$. It follows for all $v \in G \setminus \text{lin.space}(C)$. $v_k = \alpha \cdot v_i$.

Traditionally, there are two classes of reactions in a metabolic network; reversible and irreversible ones, Following our analysis, we may refine this classification and distinguish three types of reactions:

- Irreversible reactions $j \in Irr$: For all minimal proper faces G, we have either $v_i > 0$, for all $v \in G \setminus Ir$. Space (C), or $v_i = 0$, for all $v \in G$. By definition, $w_i = 0$, for all $w \in \text{lin.space}(C)$.
- 0). For all $w \in \text{lin.space}(C)$, we have again $w_i = 0$.
- Fully reversible reactions $j \in Frev$: By definition, there is $w \in lin.space(C)$ such that $w_i \neq 0$. This implies that we can find in each minimal proper face G flux vectors $v, v', v'' \in G \setminus \text{lin.space}(C)$ with $v_i > 0, v'_i < 0$ and $v''_i = 0$.

Altogether, this means that each MMB D can be characterized by a unique sign pattern P_D for the (pseudo-)irreversible reactions in the network ('+', '-', or '0'), while the flux through the fully reversible reactions may be arbitrary ('').

Example 17. In the network ILLUSNET, where Irr $\cup Rev^0 = \{1, 2, 4, 6, 7, 8\}$, the sign patterns of the MMBs D^1 , D^2 , D^3 are the following:

$$P^{1} = (+, +, \cdot, +, \cdot, 0, 0, 0, \cdot, \cdot, \cdot, \cdot)$$

$$P^{2} = (0, 0, \cdot, -, \cdot, +, +, 0, \cdot, \cdot, \cdot, \cdot)$$

$$P^3 = (0, 0, \cdot, -, \cdot, +, 0, +, \cdot, \cdot, \cdot).$$

5.2. Decomposing the network

Minimal metabolic behaviors can also be used to decompose a given metabolic network \mathcal{N} into different subnetworks. Indeed, given the sign pattern P_D of an MMB D, the set of fully reversible reactions together with the (pseudo-)irreversible reactions having a non-zero sign in P_D defines a subnetwork \mathcal{N}_D of \mathcal{N} with the following properties:

- The set of possible flux distributions in \mathcal{N}_D includes the reversible metabolic space RMS.
- All pseudo-irreversible reactions become irreversible inside \mathcal{N}_D , i.e., each pseudo-irreversible reaction operates in only one direction, which is dictated by the MMB D and given in the sign pattern P_D .
- \mathcal{N}_D is *minimal* in the sense that it includes all fully reversible reactions and a non-empty minimal set of (pseudo-)irreversible reactions that are capable of carrying flux under steady-state conditions.
- \(\mathcal{D}\)_D is functional in the sense that the set of possible flux distributions over \(\mathcal{N}_D\) includes at least one irreversible elementary flux mode.

Using the sign patterns of all the different MMBs, each corresponding to one particular minimal proper face of the flux cone, the overall metabolic network may be understood as a combination of these minimal functional subnetworks. Indeed, each possible flux distribution over the full network is a non-negative combination of possible flux distributions over the corresponding minimal functional subnetworks.

6. Computational results

In this section, we discuss how one can compute minimal metabolic behaviors, and present a number of computational results.

6.1. Computing minimal metabolic behaviors

A simple algorithm to determine the MMBs of a metabolic network is as follows. First we compute a set of generators of the flux cone C, using existing software for polyhedral computations such as cdd [5]. Note that we do not have to reconfigure the cone by splitting reversible reactions as this is done in the extreme pathway approach. If the cone C is pointed, this will be detected automatically during computation. The number of MMBs of C (which is equal to the number of extreme rays of C if C is pointed) is typically much smaller than the number of extreme rays of the reconfigured cone. Second, for each minimal proper face C of C, represented by some generator C in space C, we identify the set C of irreversible reactions C is apply the Fourier–Motzkin algorithm to eliminate the variables corresponding to the internal reversible reactions. This results in a constraint system where all internal reactions are irreversible.

If we are interested in a minimal set of generators for the flux cone C, we have to choose for each minimal proper face G^k a vector $g^k \in G^k \setminus \text{lin.space}(C)$, together with a generating set $\{b^0, \ldots, b^t\}$ of lin.space(C). If we decompose $g^k = (g^k_{\text{lir}}, g^k_{\text{Rev}^0}, g^k_{\text{Frev}})$ into components corresponding to irreversible, pseudo-irreversible, and fully reversible reactions, then the components in $(g^k_{\text{Irr}}, g^k_{\text{Rev}^0})$ are uniquely determined up to multiplication by positive scalars. There remains some freedom in the choice of the components in g^k_{Frev} . Choosing $g^k \in \text{lin.space}(C)^\perp$, i.e., $b^T g^k = 0$, for all $b \in \text{lin.space}(C)$, yields the *orthogonal representation* of the flux cone C, which is unique, but often very dense. Work in the context of the software cdd [1,15] discusses how to obtain a sparser representation of C, which is called *lexico-smallest representation*. In such a representation, the generators have a maximum set of zeroes and correspond to a subset of elementary modes.

6.2. Comparison with existing approaches

We now compare the different approaches on some example networks from the KEGG pathway database (http://www.genome.ad.jp/kegg/pathway.html). We suppose for these models that there is an unconstrained exchange flux for each metabolite that is not consumed or not produced by some internal reaction in the network. The computation of the extreme pathways, the minimal metabolic behaviors and the reversible metabolic space was done using the software cdd [5]. For computing the elementary flux modes, we used METATOOL [21]. The results are given in Tables 1 and 2.

Table 1 shows the number of internal metabolites in the network, the number of irreversible/reversible internal reactions, the number of elementary flux modes/extreme pathways/MMBs, and the dimension of the RMS. We can see that the size of our representation, given as the sum of the number of MMBs and dim(RMS), is typically much smaller than the number of extreme pathways or elementary flux modes. In various examples, the reduction is by several orders of magnitude.

Table 2 describes the distribution of the elementary flux modes and the extreme pathways inside the steady-state flux cone. We can see that a very large number of elementary flux modes and extreme pathways lies in the interior of the cone. In addition, the number of elementary flux modes/extreme pathways belonging to the minimal proper faces (see column MMB) is much larger than the number of MMBs in Table 1. This means that many elementary modes/extreme pathways belong to the same minimal proper face, which mathematically can be represented by a single vector, resp. one MMB. Similarly, the number of elementary modes/extreme pathways belonging to the reversible metabolic space is much larger than its dimension, so that there are many dependencies.

Table 1Metabolic networks, with the number of internal metabolites (Met), the number of irreversible (Irr) and reversible (Rev) internal reactions, the number of elementary flux modes (EFM), extreme pathways (EP), minimal metabolic behaviors (MMB), and the dimension of the reversible metabolic space (RMS)

Metabolic network	Met	Irr	Rev	EFM	EP	MMB	RMS
Glycolysis/Gluconeogenesis	32	18	29	19 464	1745	16	13
Citrate cycle (TCA cycle)	22	4	25	3870	1 588	4	12
Pentose phosphate pathway	34	19	24	5 155	1630	19	8
Pentose and glucuronate	50	13	46	2 258	231	7	23
Fructose and mannose	46	37	31	2 411	2 102	30	6
Galactose	41	22	28	623	524	13	9
Starch and sucrose	47	35	30	2 097	1718	30	5
Pyruvate	28	40	29	47 708	27 390	37	16
Propanoate	34	20	29	877	233	17	13
Butanoate	40	23	30	2 138	541	18	11
Nitrogen	41	53	14	601	612	44	9
Sulfur	18	26	4	321	326	28	1

Contrary to the calculation of EFMs, the calculation of EPs required a reconfiguration of the network. Except for the two-cycle extreme pathways made from a forward and a backward reaction, the set of EPs is always a subset of the set of EFMs [9].

Table 2The distribution of the elementary flux modes (EFM) and the extreme pathways (EP) in the three parts of the steady-state flux cone: the minimal proper faces (MMB), the lineality space (RMS), and the interior of the cone

Metabolic network	MMB EFM/EP	RMS EFM/EP	Interior EFM/EP
Glycolysis/Gluconeogenesis	1226/529	48/46	18190/1095
Citrate cycle (TCA cycle)	1608/568	502/258	1760/480
Pentose phosphate pathway	489/340	25/25	4641/1217
Pentose and glucuronate	1076/32	642/76	540/3
Fructose and mannose	154/148	14/14	2243/1895
Galactose	212/152	45/45	366/254
Starch and sucrose	108 / 107	8/8	1981/1565
Pyruvate	2016/1776	146/146	45 546/25 293
Propanoate	449/93	133/32	295/50
Butanoate	357/244	35/34	1746/201
Nitrogen	183/171	22/22	396/384
Sulfur	44/44	1/1	276/276

Each pair of opposite extreme pathways is considered as one reversible pathway belonging to the RMS. The two-cycle extreme pathways made from a forward and a backward reaction are not taken into account.

7. Conclusion

We have introduced a new mathematical approach to metabolic pathway analysis, characterizing a metabolic network by its minimal metabolic behaviors and the reversible metabolic space. Our method uses an outer description of the steady-state flux cone, based on sets of irreversible reactions. This is different from existing approaches, such as elementary flux modes or extreme pathways, which use an inner description, based on sets of generating vectors. The resulting description of the flux cone is minimal, unique, and satisfies a simplicity condition similar to the one that holds for elementary flux modes.

Our approach suggests a refined classification of reactions according to their reversibility type (irreversible, pseudo-irreversible, and fully reversible). While the irreversible and pseudo-irreversible reactions completely characterize minimal metabolic behaviors, the fully reversible reactions define the reversible metabolic space, which may contain useful biological information. This information is no longer explicit if we replace each reversible reaction with two irreversible ones. In our paper [11] we show that the reversibility type provides a key to elucidate reaction dependencies. Indeed, coupling relationships can only hold between reactions of a certain reversibility type. In particular, (pseudo-)irreversible reactions cannot be coupled with fully reversible reactions, and all reactions in an enzyme subset [14] must have the same reversibility type.

Finally, our description suggests a decomposition of a metabolic network into different subnetworks. Each subnetwork corresponds to one minimal proper face of the steady-state flux cone. The overall network then can be understood as a combination of these subnetworks.

Acknowledgement

The authors would like to thank Damien Eveillard for many fruitful discussions and his comments on a draft of this paper.

References

[1] D. Avis, K. Fukuda, S. Picozzi, On canonical representations of convex polyhedra, in: A. Cohen, X.S. Gao, N. Takayama (Eds.), Mathematical Software, Proceedings of the First International Congress, World Scientific, 2002, pp. 350–360.

- [2] B.L. Clarke, Stability of complex reaction networks, in: I. Prigogine, S.A. Rice (Eds.), in: Advances in Chemical Physics, vol. 43, John Wiley & Sons, 1980, pp. 1–216.
- [3] B.L. Clarke, Complete set of steady states for the general stoichiometric dynamical system, J. Chem. Phys. 75 (10) (1981) 4970-4979.
- [4] T. Dandekar, F. Moldenhauer, S. Bulik, H. Bertram, S. Schuster, A method for classifying metabolites in topological pathway analyses based on minimization of pathway number, BioSystems 70 (3) (2003) 255–270.
- [5] K. Fukuda, A. Prodon, Double description method revisited, in: Combinatorics and Computer Science, in: LNCS, vol. 1120, Springer, 1995, pp. 91–111.
- [6] J. Gagneur, S. Klamt, Computation of elementary modes: A unifying framework and the new binary approach, BMC Bioinformatics 5 (175) (2004).
- [7] R. Heinrich, S. Schuster, The Regulation of Cellular Systems, Chapman and Hall, New York, 1996.
- [8] S. Klamt, J. Stelling, Combinatorial complexity of pathway analysis in metabolic networks, Mol. Biol. Rep. 29 (1-2) (2002) 233-236.
- [9] S. Klamt, J. Stelling, Two approaches for metabolic pathway analysis? Trends Biotechnol. 21 (2003) 64-69.
- [10] S. Klamt, J. Stelling, Stoichiometric and constraint-based modeling, in: Z. Szallasi, J. Stelling, V. Périwal (Eds.), System Modeling in Cellular Biology, MIT Press, 2006, pp. 73–96 (chapter 5).
- [11] A. Larhlimi, A. Bockmayr, A new approach to flux coupling analysis of metabolic networks, in: Computational Life Sciences II, CompLife'06, Cambridge, UK, in: LNBI, vol. 4216, Springer, 2006, pp. 205–215.
- [12] B. Palsson, Systems Biology: Properties of Reconstructed Networks, Cambridge Univ. Press, 2006.
- [13] J.A. Papin, J. Stelling, N.D. Price, S. Klamt, S. Schuster, B.O. Palsson, Comparison of network-based pathway analysis methods, Trends Biotechnol. 22 (8) (2004) 400-405.
- [14] T. Pfeiffer, I. Sánchez-Valdenebro, J.C. Nuno, F. Montero, S. Schuster, Metatool: for studying metabolic networks, Bioinformatics 15 (3) (1999) 251–257.
- [15] S. Picozzi, On canonical representation of convex polyhedra, Ph.D. Thesis, Ecole Polytechnique Fédérale de Lausanne, 2005.
- [16] C.H. Schilling, D. Letscher, B. Palsson, Theory for the systemic definition of metabolic pathways and their use in interpreting metabolic function from a pathway-oriented perspective, J. Theoret. Biol. 203 (2000) 229–248.
- [17] A. Schrijver, Theory of Linear and Integer Programming, Wiley, 1986.
- [18] S. Schuster, Metabolic pathway analysis in biotechnology, in: B.N. Khodolenko, H.V. Westerhoff (Eds.), Metabolic Engineering in the Post Genomic Era, MIT Press, 2004, pp. 181–208 (chapter 8).
- [19] S. Schuster, C. Hilgetag, On elementary flux modes in biochemical systems at steady state, J. Biol. Syst. 2 (1994) 165-182.
- [20] S. Schuster, C. Hilgetag, J.H. Woods, D.A. Fell, Reaction routes in biochemical reaction systems: Algebraic properties, validated calculation procedure and example from nucleotide metabolism, J. Math. Biol. 45 (2002) 153–181.
- [21] A. von Kamp, S. Schuster, Metatool 5.0: fast and flexible elementary modes analysis, Bioinformatics 22 (15) (2006) 1930-1931.
- [22] C. Wagner, R. Urbanczik, The geometry of the flux cone of a metabolic network, Biophys. J. 89 (2005) 3837–3845.