

ON ELEMENTARY FLUX MODES IN BIOCHEMICAL REACTION SYSTEMS AT STEADY STATE

STEFAN SCHUSTER and CLAUS HILGETAG

*Fachbereich Biologie, Institut für Biophysik,
Humboldt-Universität zu Berlin, Invalidenstr. 42, D-10115 Berlin, Germany*

Received 11 August 1993

Revised 3 January 1994

ABSTRACT

A mathematical definition of the concept of elementary mode is given so as to apply to biochemical reaction systems subsisting at steady state. This definition relates to existing concepts of null-space vectors and includes a condition of simplicity. It is shown that for systems in which all fluxes have fixed signs, all elementary modes are given by the generating vectors of a convex cone and can, thus, be computed by an existing algorithm. The present analysis allows for the more general case that some reactions can proceed in either direction. Basic ideas on how to compute the complete set of elementary modes in this situation are outlined and verified by way of several examples, with one of them representing glycolysis and gluconeogenesis. These examples show that the elementary modes can be interpreted in terms of the particular biochemical functions of the network. The relationships to (futile) substrate cycles are elucidated.

Keywords: Biochemical pathways, mathematical modelling, elementary mode, substrate cycle.

1. Introduction

Investigation of steady states plays an important role in the modelling of biochemical reaction systems, because virtually stationary regimes are frequently encountered in experimental settings and under *in-vivo* conditions [3,9,16]. As for systems with oscillatory behaviour, stationary concentrations can be calculated as average values over a longer time span. In both situations, reaction rates have to fulfill balance equations since for each substance of the reaction network, (averaged) inputs have to balance (averaged) outputs. In mathematical terms, this means that the vector of stationary reaction rates, called fluxes below, has to be situated in the null-space (kernel) of the stoichiometry matrix, N [11,16]. This null-space can be represented as being spanned by basis vectors. It is of interest to find those basis vectors that can be interpreted in terms of biochemical functioning. In particular, different functions, such as ATP production and amino acid synthesis, should be represented by separate vectors. So one attempts, guided by the principle of Ockham's razor, to

find the simplest biochemically meaningful flux vectors possible, in terms of which all other flux vectors can be expressed.

Leiser and Blum [13] put forward the idea to define “fundamental modes” of substrate cycles by invoking that any steady-state flux pattern could be decomposed as a linear superposition of these modes and that these modes are all “realizable”, i.e., no thermodynamic laws or hypotheses about enzyme irreversibility should be violated. Sign constraints for fluxes are frequent in modelling studies, in particular if the reaction rates are defined as unidirectional rates [1,2]. This is essential, for example, in the description of tracer dynamics, where forward and reverse reactions enter the equations in a different way [20]. As for virtually irreversible reactions, also the net rates can be assumed to be non-negative [9,12].

Fell [5,7] proposed to define fundamental modes by a proper choice of basis vectors of the null-space. He observed that this method meets the difficulties that irreversibility constraints may be violated and that there may be a greater number of relevant ways on how to connect the inputs to the outputs of the system than vectors are needed to form a basis.

Studying reaction networks with all of their fluxes subject to sign restriction and using methods of convex analysis, we defined the fundamental modes as the generating vectors of the cone of admissible fluxes, and presented an algorithm for computing these generating vectors [19].

In the present paper, we expand that analysis by allowing for the situation that some reactions proceed in a fixed direction whereas others may proceed in either direction. Fluxes that are not restricted with respect to their sign, have frequently to be considered in the modelling of biochemical systems (cf. also [13]). This concerns, in particular, the fluxes through reactions shared by catabolic and anabolic pathways, e.g., glycolysis and gluconeogenesis, and fluxes through the reversible reactions of the hexose monophosphate shunt.

In Sec. 3, we shall formalize the concept of fundamental mode (as explained below, we prefer the term “elementary mode”). In Sec. 4, we shall study the situation that all fluxes have fixed signs and explain how then to compute the elementary modes. The case that some fluxes are not restricted to one direction will be treated in Sec. 5. We shall present some basic ideas on how to compute the elementary modes in that case. Section 6 illustrates the analysis by way of biochemical examples.

2. The Space of All Mathematically Admissible Steady-State Vectors

We follow the formalism of describing the behaviour of biochemical systems by ordinary differential equations [6,9,11,16],

$$\frac{d\mathbf{X}}{dt} = \mathbf{N}\mathbf{V}(\mathbf{X}) \quad (2.1)$$

where \mathbf{N} is the stoichiometry matrix of the system, \mathbf{V} denotes the vector of reaction rates and \mathbf{X} stands for the vector of concentrations. The dimension of \mathbf{N} is given

by the number of (internal) reacting species, n , and the number of reactions, r . Steady-state conditions are defined by the algebraic equation system

$$\mathbf{N}\mathbf{V} = \mathbf{0}, \quad (2.2)$$

which results from Eq. (2.1).

Non-trivial solutions to this system only exist if there are linear dependencies between the columns of \mathbf{N} , i.e., if

$$\nu = \text{rank}(\mathbf{N}) \quad (2.3)$$

is less than r . These dependencies can be expressed by the null-space matrix, \mathbf{K} ,

$$\mathbf{N}\mathbf{K} = \mathbf{0}. \quad (2.4)$$

Accordingly, the columns of \mathbf{K} span the subspace of all steady-state vectors mathematically compatible with the stoichiometric structure of the system. In addition, the values of fluxes may further be restricted by the rate laws, $\mathbf{V}(\mathbf{X})$. This issue will not be dealt with in the present paper. We follow here an idea put forward by several authors [2,16] saying that the theoretical investigation of reaction systems should begin with an analysis of the structural properties (the invariants) of the systems, since the kinetic properties are subject to frequent changes (e.g., by inhibition and activation of enzymes).

Equation (2.4) shows that \mathbf{K} is indeterminate to the extent of postmultiplication by a non-singular square matrix. One is therefore interested in finding an appropriate, "canonical", representation of the null-space matrix. In the computer program "CONTROL" [14], \mathbf{K} is computed to have the form

$$\mathbf{K} = \begin{pmatrix} \mathbf{K}_0 \\ \mathbf{I} \end{pmatrix}, \quad (2.5)$$

where \mathbf{I} denotes the $(r - \nu) \times (r - \nu)$ identity matrix. This can be achieved, for example, by using the Gaussian elimination method. Since this representation has the advantage that it contains a considerable number of zeros, it meets, to a certain extent, the criterion of simplicity of the particular vectors. Importantly, the representation (2.5) transforms into a block-diagonal form by permutation of columns and rows, provided \mathbf{K} is block-diagonalizable [22]. The blocks of \mathbf{K} correspond to subnetworks the fluxes of which are completely independent.

Some problems involved in using basis null-space vectors as fundamental modes reveal already here. Since it is not uniquely determined which $r - \nu$ columns of \mathbf{N} are chosen so as to correspond to the submatrix \mathbf{I} , there is still some ambiguity in calculating matrix \mathbf{K} according to Eq. (2.5). Furthermore, on calculating \mathbf{K} , it may occur that elements of \mathbf{K}_0 are negative although the corresponding flux should be non-negative for thermodynamic reasons.

3. Definition of Elementary Modes

On studying biochemical reaction systems, one can generally distinguish between boundary reactions and internal reactions. Boundary reactions have the property

that all the reactants utilized or all the products formed by these reactions are external (i.e. their concentrations are fixed). In the example shown in Fig. 1, reactions 1, 2 and 5 are boundary reactions, while reactions 3 and 4 are internal. This distinction can be made by investigating the stoichiometry matrix. Boundary reactions are represented by those columns that only contain elements of the same sign or zeros, while columns corresponding to internal reactions contain elements of opposite sign, as can be exemplified by the stoichiometry matrix belonging to the system shown in Fig. 1,

$$\mathbf{N} = \begin{pmatrix} 1 & -1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 & -1 \end{pmatrix}. \quad (3.1)$$

Although the stoichiometry matrix does not contain any information about external metabolites, it can at least be said that for all reactions that have no internal metabolites at one side of the reaction arrow, external reactants must be located there.

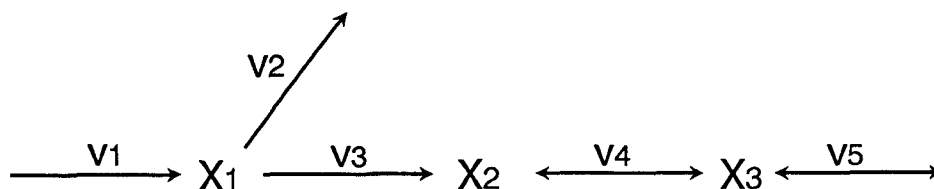


Fig. 1. Branched reaction system.

Assume now that reactions 1, 2, and 3 in this system are irreversible in the direction indicated by the arrows. It is then plausible to attach to the system two elementary modes given by the flux vectors

$$\mathbf{V}^{(1)} = (1 \ 1 \ 0 \ 0 \ 0)^T, \quad \mathbf{V}^{(2)} = (1 \ 0 \ 1 \ 1 \ 1)^T, \quad (3.2a,b)$$

because each vector of steady-state fluxes not violating the irreversibility conditions can be written as non-negative linear combination of $\mathbf{V}^{(1)}$ and $\mathbf{V}^{(2)}$.

It seems natural that the term “elementary mode” should be so general that two vectors that differ only by a positive factor are not considered different. For the above example, this means that we may choose, instead of the vectors given in Eq. (3.2), $\mathbf{V}^{(1)} = (2 \ 2 \ 0 \ 0 \ 0)^T$, $\mathbf{V}^{(2)} = (3 \ 0 \ 3 \ 3 \ 3)^T$. Thus, elementary modes are not simply flux vectors, but classes of vectors that convert into one another by multiplication by a positive scalar. Two elementary modes that differ by a negative factor are to be considered different, because opposite directions of flux correspond to different biological functions, e.g. the ATP production catalysed by the H^+ -ATPase in the mitochondrial inner membrane and the proton transport catalysed by the H^+ -ATPase in plant vacuolar membranes.

The irreversibility constraint can always be written as a non-negativity condition, since in case of negative flux we can reverse the orientation of reaction without loss of generality. By decomposing the flux vector into two subvectors \mathbf{V}^{irr} and \mathbf{V}^{rev} , we can express this condition as

$$\mathbf{V}^{\text{irr}} \geq 0. \quad (3.3)$$

Here, \mathbf{V}^{irr} represents reaction rates which for thermodynamic or other reasons have to be non-negative, while \mathbf{V}^{rev} stands for the fluxes that are not subject to sign restriction. Lacking a better notion, we call the reactions represented by \mathbf{V}^{rev} reversible reactions, although those reactions that have a fixed sign of flux for some of the reasons mentioned in the Introduction may be reversible in a physico-chemical sense. By “reversible reactions”, we here mean reactions that can proceed in either direction under physiological conditions, such as the reactions shared by glycolysis and gluconeogenesis.

The criterion that elementary modes be the simplest flux vectors possible cannot be phrased by just invoking that these vectors have to contain a maximum number of zeros each. If this were done for the system shown in Fig. 1, only $\mathbf{V}^{(1)}$ could represent an elementary mode. Thus, the principle of simplicity has rather to be phrased as a criterion about a “relative” maximum of zeros, i.e., there should be no flux vectors that have more zeros and give rise to the vector under consideration by non-negative combination.

The above reasoning leads us to

Definition 3.1:

A flux mode, \mathbf{M} , is defined as the set

$$\mathbf{M} = \{\mathbf{V} \in \mathbb{R}^r \mid \mathbf{V} = \lambda \mathbf{V}^*, \lambda > 0\}, \quad (3.4)$$

where \mathbf{V}^* is an r -vector (unequal to the null vector) fulfilling the following two conditions,

(C1) Steady-state condition. \mathbf{V}^* satisfies Eq. (2.2).

(C2) Sign restriction. \mathbf{V}^* contains a subvector, \mathbf{V}^{irr} , that fulfills inequality (3.3), with the indices of the components of \mathbf{V}^{irr} being the indices of those fluxes that are subject to a non-negativity condition.

Remark:

From this definition, it follows that \mathbf{V}^* is itself a member of the set \mathbf{M} . It can be taken as a representative of \mathbf{M} . Moreover, it is obvious that besides \mathbf{V}^* , every element of \mathbf{M} obeys conditions (C1) and (C2).

Definition 3.2:

A flux mode \mathbf{M} with a representative \mathbf{V}^* is called elementary (flux) mode if, and only if, \mathbf{V}^* fulfills the condition

(C3) Simplicity. There exists no couple of vectors \mathbf{V}' , \mathbf{V}'' (unequal to the null vector) with the following properties:

- (i) \mathbf{V}^* is a non-negative linear combination of \mathbf{V}' and \mathbf{V}'' ,

$$\mathbf{V}^* = \lambda_1 \mathbf{V}' + \lambda_2 \mathbf{V}'', \quad \lambda_1, \lambda_2 > 0. \quad (3.5)$$

- (ii) \mathbf{V}' , \mathbf{V}'' obey restrictions (C1) and (C2),

- (iii) both \mathbf{V}' and \mathbf{V}'' contain at least the same number of zero elements as \mathbf{V}^* , and at least one of them contains more zero elements than \mathbf{V}^* ,

- (iv) for all indices i corresponding to boundary reactions, the components are not of opposite sign, $\text{sgn}(\nu'_i) \neq -\text{sgn}(\nu''_i)$.

It is worth noting that due to the non-recursive structure of these definitions, the elementary modes of any system are uniquely determined. Moreover, there is some similarity between Definition 3.1 and the concept of quotient space. If no sign restriction applied, the flux modes were elements of a suitably defined quotient space.

Definition 3.3:

A flux mode \mathbf{M} is called reversible flux mode if, and only if, $\mathbf{M}' = \{-\mathbf{V} | \mathbf{V} \in \mathbf{M}\}$ is a flux mode as well. Otherwise, \mathbf{M} is called irreversible flux mode.

The same distinction can then be made for elementary modes.

Definition 3.4:

A flux mode \mathbf{M} with a representative \mathbf{V}^* that does not contain any components corresponding to boundary reactions is termed cyclic mode (or cycle). An elementary cyclic mode (or elementary cycle) is a cyclic mode fulfilling condition (C3).

The latter definition is based on an idea put forward by Fell [5,7] saying that substrate cycles should not involve any reactions that connect to the external pools.

If we assume that in the example shown in Fig. 1 only reaction 1 is irreversible, we obtain the elementary flux vectors $(1 \ 1 \ 0 \ 0 \ 0)^T$, $(1 \ 0 \ 1 \ 1 \ 1)^T$, $(0 \ -1 \ 1 \ 1 \ 1)^T$ and $(0 \ 1 \ -1 \ -1 \ -1)^T$. The former two represent irreversible modes and the latter two, reversible modes.

The reaction system shown in Fig. 2 can serve for illustration of condition (C3). The vector $(0 \ 1 \ -1 \ 1 \ 0 \ 1)^T$ is no elementary mode since it is the sum of $(0 \ 1 \ -1 \ 0 \ -1 \ 0)^T$ and $(0 \ 0 \ 0 \ 1 \ 1 \ 1)^T$, both of which have more zeros than the first vector and have the same sign or zeros at the positions related to boundary reactions (ν_1 , ν_2 and ν_3).

Now consider the branched system depicted in Fig. 3. Here, the vector $(1 \ 1 \ 1 \ 0 \ 0)^T$ is an elementary mode, although it is the sum of $(1 \ 0 \ 0 \ 1 \ 0)^T$ and $(0 \ 1 \ 1 \ -1 \ 0)^T$. Since the latter two vectors have opposite signs at position 4, which corresponds to a boundary reaction, they do not meet condition (C3, iv). We deliberately imposed this condition in order not to exclude potential elementary modes representing main pathways, even if they are longer than other parts of the system.

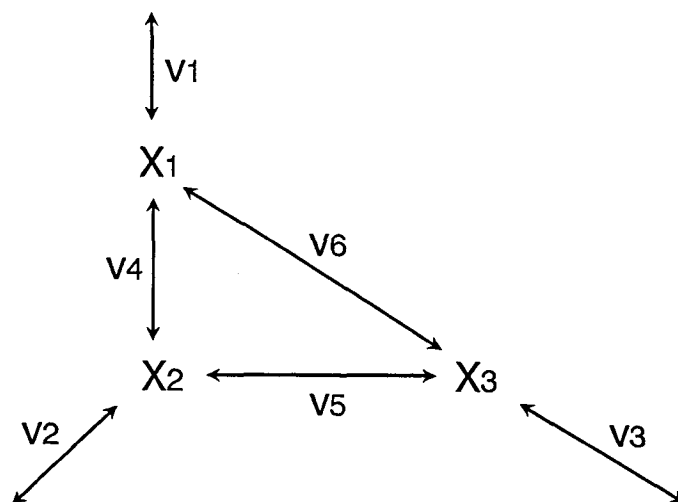


Fig. 2. Cyclic reaction system. This scheme can stand, for example, for the PYR/OAA/PEP cycle and was also studied in [13].

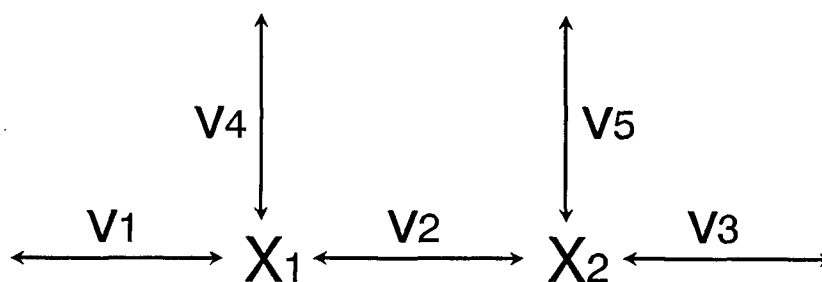


Fig. 3. System with two ramification points.

Although every elementary mode that is not cyclic has obviously to involve at least one input and one output flux, there exist systems with elementary modes that include only one boundary reaction, because also internal reactions may connect to external metabolites. An example is shown in Fig. 4. It has the stoichiometry matrix

$$N = \begin{pmatrix} 1 & -2 & 1 \\ 0 & 1 & -1 \end{pmatrix}. \quad (3.6)$$

If all reactions are irreversible, the only elementary mode is $(1 \ 1 \ 1)^T$.

To elucidate the mathematical implications of Definitions 3.1 and 3.2, we first deal with the question as to what region in the flux space is formed by the flux vectors fulfilling relations (2.2) and (3.3). In convex analysis, it is shown that the region determined by a linear homogeneous equation/inequality system is a convex polyhedral cone, \mathcal{C} [15,18]. Convex analysis further states that every point of such a

cone is a non-negative combination of fundamental vectors, $\mathbf{f}^{(k)}$, and basis vectors, $\mathbf{b}^{(m)}$,

$$\mathcal{C} = \left\{ \mathbf{V} : \mathbf{V} = \sum_{k=1}^p \eta_k \mathbf{f}^{(k)} + \sum_{m=1}^s \lambda_m \mathbf{b}^{(m)}, \eta_k, \lambda_m \geq 0 \quad \forall k, m \right\}. \quad (3.7)$$

Both fundamental and basis vectors are also called generating vectors or extreme rays. There are minimum numbers of fundamental and basis vectors that are sufficient to span the cone. The basis vectors, $\mathbf{b}^{(m)}$, are those extreme rays of cone \mathcal{C} for which also the negative vector, $-\mathbf{b}^{(m)}$, is contained in \mathcal{C} .

Usage of the terms basis and fundamental vectors in convex analysis [15] is the reason why we prefer the term elementary mode rather than fundamental mode, as used in [13].

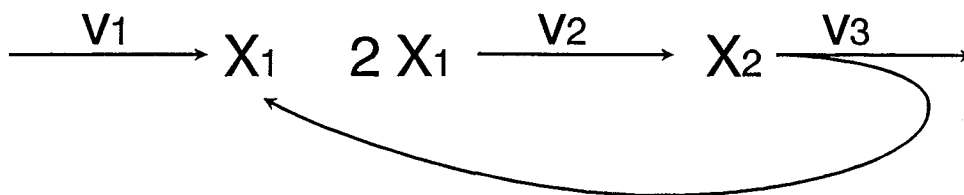


Fig. 4. System with only one boundary reaction.

4. All Reactions are Considered Irreversible

In a previous paper [19], we studied the situation that all fluxes of the system are subject to non-negativity conditions, i.e.,

$$\mathbf{V} \geq 0. \quad (4.1)$$

Let \mathcal{K} denote the cone determined by relations (2.2) and (4.1). \mathcal{K} is a pointed cone, that is, any two vectors contained in the cone make an angle of less than, or equal to, 90 degrees. Relation (4.1) implies that \mathcal{K} has no basis vectors. The fundamental vectors can be determined by an algorithm given in [19]. Generating vectors of pointed convex polyhedral cones are unique up to multiplication by positive scalars [18]. It is worth mentioning that cone \mathcal{K} can have any dimension from zero to $r - \nu$, depending on how the null-space is situated relative to the positive orthant. Furthermore, the number of generating vectors may be greater than the dimension of the cone.

An important relationship is expressed in the following

Theorem 4.1:

The generating vectors of the cone \mathcal{K} determined by Eq. (2.2) and inequality (4.1) constitute a complete set of representatives of the elementary modes under the sign restriction (4.1).

Proof:

All generating vectors fulfill, by definition, conditions (C1) and (C2). Since generating vectors of pointed convex cones are those vectors of the cone that cannot be written as non-negative linear combination of other vectors belonging to the cone [18], they fulfill condition (C3). Hence, all generating vectors are representatives of elementary modes.

Assume now that there is a vector \mathbf{V} fulfilling conditions (C1) to (C3) that is no generating vector of \mathcal{K} . \mathbf{V} would then be a non-negative linear combination of at least two different generating vectors. It can then also be written as a combination of exactly two vectors, one of them being a generating vector,

$$\mathbf{V} = \sum_{k=1}^p \lambda_k \mathbf{f}^{(k)} = \lambda_1 \mathbf{f}^{(1)} + \sum_{k=2}^p \lambda_k \mathbf{f}^{(k)}, \quad \lambda_1 \neq 0, \quad (4.2)$$

As we here assume all reactions to be irreversible, any two vectors of \mathcal{K} cannot have opposite signs at the same position, so that the two vectors $\mathbf{f}^{(1)}$ and $\sum_{k=2}^p \lambda_k \mathbf{f}^{(k)}$ meet condition (iv) of (C3). Furthermore, $\mathbf{f}^{(1)}$ has more zeros than \mathbf{V} because the generating vectors have the property that a maximum number of inequality side constraints are fulfilled as equality, which means in our case that a maximum number of fluxes are zero. Therefore, \mathbf{V} can be expressed into two vectors satisfying conditions (i) to (iv). This leads to the contradiction that \mathbf{V} does not fulfil condition (C3), which completes the proof.

Due to the above Theorem, the algorithm presented in [19] can serve to compute all elementary modes.

5. Some Reactions are Considered Reversible

We can distinguish the three following cases:

- The system has only irreversible elementary modes, although some reactions of the system are considered reversible. For example, in the system shown in Fig. 1 with all reactions treated reversible except reactions 1 and 3, no reversible elementary mode occurs.
- There are irreversible as well as reversible elementary modes. Consider the reaction system shown in Fig. 5. It has the reversible elementary modes

$$\mathbf{b}^{(1)} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad \mathbf{b}^{(2)} = \begin{pmatrix} -1 \\ 0 \\ -1 \end{pmatrix}, \quad (5.1a,b)$$

and the irreversible elementary modes

$$\mathbf{f}^{(1)} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{f}^{(2)} = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}. \quad (5.2a,b)$$

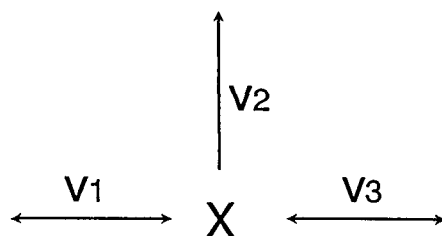


Fig. 5. Simple branched system with one irreversible reaction.

A representative of one irreversible elementary mode together with the reversible elementary modes would, however, be sufficient to span cone \mathcal{C} , as seen in Fig. 6. Thus, the example shows that for systems containing some reversible reactions, there may be a greater number of elementary modes than generating vectors are needed to span cone \mathcal{C} , according to Eq. (3.7). It can further be observed that choice of the fundamental vector is now not unique; we can choose $\mathbf{f}^{(1)}$, $\mathbf{f}^{(2)}$, or any non-negative linear combination of these, and of the basis vectors, e.g., the vector $\mathbf{f}^* = (-1 \ 2 \ 1)^T$, which is orthogonal to the basis vectors.

In this particular example, the basis vectors are unique (up to multiplication by positive scalars), since there is only one pair of them. If more basis vectors exist,

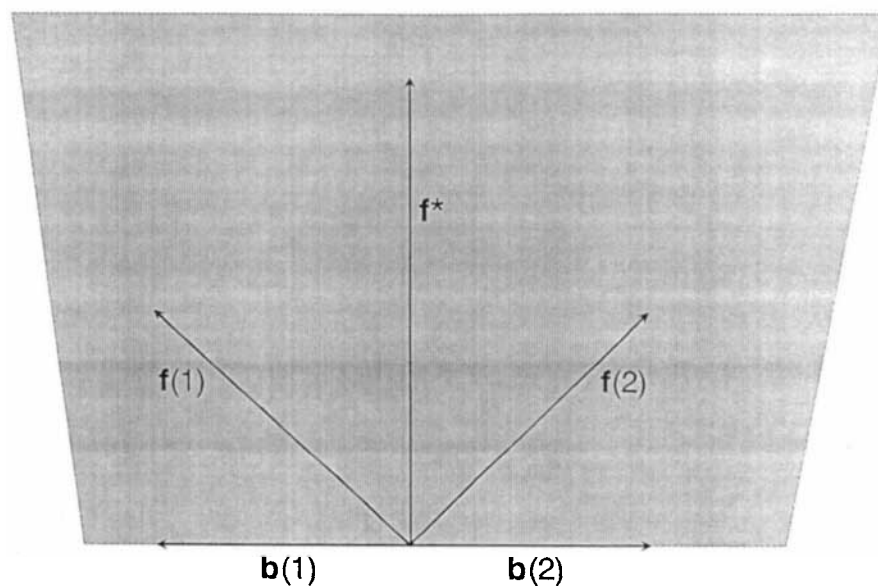


Fig. 6. Cone of admissible steady state fluxes for the system shown in Fig. 5. Notations: $\mathbf{b}(1)$ and $\mathbf{b}(2)$, basis vectors; \mathbf{f}^* , fundamental vector orthogonal to the basis vectors; $\mathbf{f}(1)$ and $\mathbf{f}(2)$, fundamental vectors representing elementary modes.

they span a sub-space of the kernel of the stoichiometry matrix with a dimension greater than one and are hence not unique (see Sec. 2).

- (c) There are only reversible elementary modes (and only basis vectors). This can also occur if some reactions are irreversible, namely if the respective fluxes are zero. Importantly, the number of reversible elementary modes need not simply relate to the number of basis vectors, as seen in the system shown in Fig. 3, which has $2 \cdot 6 = 12$ reversible elementary modes, while $r - \nu = 3$.

Whereas in the situation that all reactions are irreversible, all generating vectors of cone \mathcal{K} (which are representatives of the elementary modes) are edges of this cone, this need not be so in the situation now considered. For example, in the system shown in Fig. 5, we cannot find any fundamental vector lying on an edge of cone \mathcal{C} (see Fig. 6). Conversely, however, every edge of \mathcal{C} corresponds to an elementary mode because it fulfills conditions (C1) and (C2) and, since it cannot be expressed as a non-negative linear combination of other vectors of the cone, also condition (C3).

An algorithm for detecting the elementary modes of systems containing reversible reactions can be developed on the basis of an algorithm for computing the generating vectors of convex polyhedral cones [15]. It is also related to the algorithm presented previously [19,21] for systems containing irreversible reactions only. In that situation, due to condition (4.1), only fundamental vectors had to be dealt with, which were obtained by a step-wise calculation of tableaux, $\mathbf{T}^{(j)}$. In the situation of reversible reactions, we start from a tableau containing the transposed stoichiometry matrix and an identity matrix of dimension $r \times r$,

$$\mathbf{T}^{(0)} = \begin{pmatrix} \mathbf{B}^{(0)} \\ \mathbf{F}^{(0)} \end{pmatrix} = \begin{pmatrix} \mathbf{I} & \mathbf{0} & \mathbf{N}_{\text{rev}}^T \\ \mathbf{0} & \mathbf{I} & \mathbf{N}_{\text{irr}}^T \end{pmatrix}, \quad (5.3)$$

where the decomposition of \mathbf{N} into \mathbf{N}_{rev} and \mathbf{N}_{irr} is done according to the decomposition of \mathbf{V} into \mathbf{V}^{rev} and \mathbf{V}^{irr} .

If all metabolites were external ones, we would not need consider the steady-state condition (2.2). Due to the non-negativity condition (3.3), the row vectors of $(\mathbf{I} \ \mathbf{0})$ would then be a complete set of representatives of irreversible elementary modes, and the row vectors of $(\mathbf{0} \ \mathbf{I})$ together with the row vectors of $(\mathbf{0} \ -\mathbf{I})$ may be taken as representatives of the reversible elementary modes. For the sake of simplicity, any two preliminary reversible elementary modes $\mathbf{b}^{(k)}$ and $-\mathbf{b}^{(k)}$ can, in the algorithm, be replaced by only one of them, and at the end, the set of reversible elementary modes is enlarged by including the opposites of all the ones calculated. Thus, the rows of the identity matrix in Eq. (5.3) represent the irreversible and reversible elementary modes of the system with all metabolites considered external. Now, one successively considers the particular equations contained in the matrix equation (2.2). Condition (3.3) together with the first j equations contained in Eq. (2.2) determines a convex polyhedral cone, \mathcal{C}_j . Upon including a further, $(j+1)$ st equation out of Eq. (2.2), a cone, \mathcal{C}_{j+1} , obtains, which is a subset

of \mathcal{C}_j . Each of its generating vectors is either a generating vector of \mathcal{C}_j as well, or it can be written as non-negative linear combination of two generating vectors of \mathcal{C}_j , because the extreme vectors of \mathcal{C}_{j+1} are determined as the intersection of the $(r-1)$ -dimensional hyperplane given by the $(j+1)$ st equation out of Eq. (2.2) and a two-dimensional face of \mathcal{C}_j . As pointed out in Sec. 4, there may be elementary modes that are not represented by edges of the cone, but lie within the cone. Due to the condition of simplicity, and by analogy to the properties of generating vectors, one can assume that also those elementary modes are non-negative linear combinations of two elementary modes of cone \mathcal{C}_j each. Therefore, in each step of the algorithm, rows of the $(j+1)$ st tableau obtain as non-negative combination of two rows each of the j th tableau representing \mathcal{C}_j , so that the $(j+1)$ st column of \mathbf{N}^T becomes the null vector. In extension to the algorithm presented in [19], now also preliminary reversible elementary modes have to be combined to give such modes in $\mathbf{T}^{(j+1)}$,

$$\mathbf{b}^* = t_{m,r+j+1}^{(j)} \cdot \mathbf{t}_i^{(j)} - t_{i,r+j+1}^{(j)} \cdot \mathbf{t}_m^{(j)}, \quad \mathbf{t}_i^{(j)}, \quad \mathbf{t}_m^{(j)} \in \mathbf{B}^{(j)}, \quad i \neq m, \quad (5.4)$$

where $\mathbf{t}_i^{(j)}$ and $\mathbf{t}_m^{(j)}$ denote the i th and m th row vectors of $\mathbf{T}^{(j)}$.

In order that these vectors \mathbf{b}^* are not equal to a combination of two other elementary vectors obtained in the same step, a similar condition as in combining irreversible modes has to be included (Eq. (14) in [19]). Moreover, some reversible elementary modes have also to be combined with irreversible elementary modes, to give fundamental vectors in $\mathbf{T}^{(j+1)}$, additional to the ones found by the algorithm presented in [19],

$$\begin{aligned} \mathbf{f}^* &= \text{sgn } t_{m,r+j+1}^{(j)} (t_{m,r+j+1}^{(j)} \cdot \mathbf{t}_i^{(j)} - t_{i,r+j+1}^{(j)} \cdot \mathbf{t}_m^{(j)}), \\ \mathbf{t}_m^{(j)} &\in \mathbf{B}^{(j)}, \quad \mathbf{t}_i^{(j)} \in \mathbf{F}^{(j)}, \quad t_{m,r+j+1}^{(j)} \neq 0. \end{aligned} \quad (5.5)$$

The algorithm ends with tableau $\mathbf{T}^{(n)}$. After including the submatrix $-\mathbf{B}^{(n)}$, the submatrix of $\mathbf{T}^{(n)}$ consisting of the r left-hand side columns contains, as rows, the elementary modes.

In spite of the relatively simple definition of elementary modes, a general algorithm for computing the complete set of these modes, as outlined afore, turns out to be rather intricate, and requires elaborate mathematical reasoning. Therefore, we leave the detailed description and justification of the algorithm to a future publication.

6. Biochemical Examples

We illustrate the algorithm outlined in the previous section by way of the example shown in Fig. 2, which may serve as a model of the phosphoenolpyruvate/pyruvate/oxaloacetate cycle [13]. Here, we assume the reactions 4, 5, and 6 to be irreversible so as to operate in clockwise orientation. The starting tableau reads (zeros are omitted for clarity's sake).

$$\mathbf{T}^{(0)} = \left(\begin{array}{ccccccccc} 1 & & & & & \vdots & & 1 & \\ & 1 & & & & \vdots & & & 1 \\ & & 1 & & & \vdots & & & 1 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ & & & 1 & & \vdots & 1 & -1 & \\ & & & & 1 & \vdots & & 1 & -1 \\ & & & & & 1 & -1 & 1 & \\ \mathbf{I} & & & & & & \mathbf{N}^T & & \end{array} \right) \begin{array}{l} \mathbf{B}^{(0)} \\ \\ \\ \\ \mathbf{F}^{(0)} \end{array} \quad (6.1)$$

We first examine the left-hand side column of $\mathbf{N}^T \cdot \mathbf{F}^{(1)}$ results from summation of the fourth and sixth row, because the elements at positions (4,7) and (6,7) are non-zero and of opposite sign, and from combination of the first row with all rows of $\mathbf{F}^{(0)}$, according to formula (5.5). $\mathbf{B}^{(1)}$ obtains by combining the first row with the second and third rows according to formula (5.4). This gives

$$\mathbf{T}^{(1)} = \left(\begin{array}{ccccccccc} & 1 & & & & \vdots & & 1 & \\ & & 1 & & & \vdots & & & 1 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ -1 & & & 1 & & \vdots & & -1 & \\ & & & & 1 & \vdots & & 1 & -1 \\ 1 & & & & & 1 & \vdots & & 1 \\ & & & 1 & & 1 & \vdots & -1 & 1 \end{array} \right) \quad (6.2)$$

In a similar way, we obtain

$$\mathbf{T}^{(2)} = \left(\begin{array}{ccccccccc} & & 1 & & & \vdots & 0 & 0 & 1 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ -1 & 1 & & 1 & & \vdots & & & \\ & -1 & & & 1 & \vdots & & -1 & \\ 1 & & & & & 1 & \vdots & & 1 \\ & 1 & & 1 & & 1 & \vdots & & 1 \\ -1 & & 1 & 1 & & \vdots & & -1 & \\ & & 1 & 1 & 1 & \vdots & & & \end{array} \right) \quad (6.3)$$

Since $\mathbf{B}^{(2)}$ contains one vector only, which has a non-zero element at the last position, $\mathbf{B}^{(3)}$ is empty. Thus, we only obtain irreversible elementary modes,

$$\mathbf{T}^{(3)} = \mathbf{F}^{(3)} = \begin{pmatrix} -1 & 1 & 0 & 1 & 0 & 0 & \vdots & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 1 & 0 & \vdots & 0 & 0 & 0 \\ 1 & 0 & -1 & 0 & 0 & 1 & \vdots & 0 & 0 & 0 \\ 0 & 1 & -1 & 1 & 0 & 1 & \vdots & 0 & 0 & 0 \\ -1 & 0 & 1 & 1 & 1 & 0 & \vdots & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & \vdots & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & 1 & 1 & \vdots & 0 & 0 & 0 \end{pmatrix} = (\mathbf{f}^{(k)} : \mathbf{0}). \quad (6.4)$$

The combinations of the third and fifth row of $\mathbf{T}^{(2)}$, and of the fourth and sixth row of $\mathbf{T}^{(2)}$, and of the fifth and sixth row of $\mathbf{T}^{(2)}$ have to be ruled out by a condition imposed on the combination of fundamental vectors (Eq. (14) in [19]).

Figure 7 illustrates the seven (irreversible) elementary modes. $\mathbf{f}^{(1)}$, for example, represents the elementary mode $-\nu_1 = \nu_2 = \nu_4$, $\nu_3 = \nu_5 = \nu_6 = 0$. The fact that no reversible elementary mode is obtained, corresponds to the feature that due to the irreversibility of reactions 4, 5, and 6, no elementary mode of the system shown in Fig. 2 can be inverted to obtain another elementary mode. $\mathbf{f}^{(1)}$, $\mathbf{f}^{(2)}$ and $\mathbf{f}^{(3)}$ form a complete set of generating vectors of cone \mathcal{C} . Although the remaining modes contained in $\mathbf{T}^{(3)}$ can be obtained by non-negative linear combination of generating vectors, they are elementary modes since they fulfill condition (C3). Although the cyclic mode $(0 \ 0 \ 0 \ 1 \ 1 \ 1)^T$ does not belong to the generating vectors, it contains as many zeros as these. The fact that it obtains as an elementary mode is in support of the appropriateness of condition (C3). The remaining three elementary modes encompass one zero less, but they are elementary in the sense that there is no route of the same orientation that connects the same two external metabolites in a simpler way. For instance, the mode $(0 \ 1 \ -1 \ 1 \ 0 \ 1)^T$ would also obtain as the sum of $\mathbf{f}^{(1)}$ and $\mathbf{f}^{(3)}$, but this would involve an extra boundary reaction and is therefore ruled out by condition (C3, iv). The same external metabolites are also connected by the mode $\mathbf{f}^{(2)}$, which is oriented, however, in the opposite direction and is, hence, not equivalent in terms of biochemical functioning.

Leiser and Blum [13] also indicate seven fundamental modes for this system, with one of them being the (futile) cyclic mode, and the other six being non-futile. As these modes have five non-zero components each, they do not, to our eyes, fulfill a condition of simplicity.

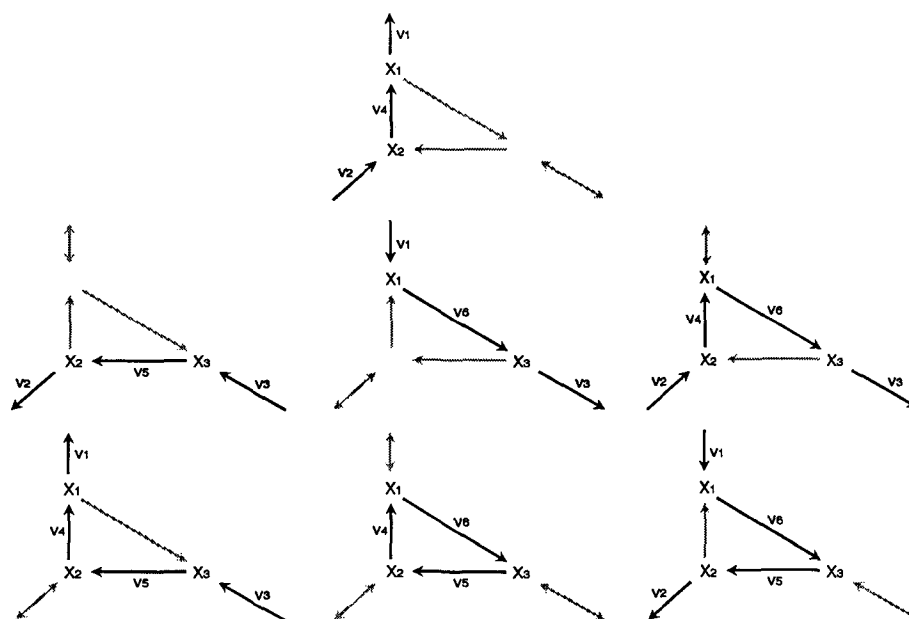


Fig. 7. Elementary modes of the PYR/OAA/PEP cycle. In contrast to Fig. 2, the reactions within the cycle are here considered irreversible.

If all reactions in the considered system are assumed to be reversible, as shown in Fig. 2, one obtains reversible elementary modes only, as given by

$$\mathbf{b}^{(1)} = (-1 \ 1 \ 0 \ 1 \ 0 \ 0)^T, \quad \mathbf{b}^{(2)} = (0 \ -1 \ 1 \ 0 \ 1 \ 0)^T, \quad (6.5a,b)$$

$$\mathbf{b}^{(3)} = (1 \ 0 \ -1 \ 0 \ 0 \ 1)^T, \quad \mathbf{b}^{(4)} = (0 \ 0 \ 0 \ 1 \ 1 \ 1)^T, \quad (6.5c,d)$$

and $-\mathbf{b}^{(k)}$, $k = 1, \dots, 4$, where, for instance, $\mathbf{b}^{(1)}$, $\mathbf{b}^{(2)}$ and $\mathbf{b}^{(3)}$ can be taken as a complete set of generating vectors. Now, modes with less than three zero elements drop out because they can be written as the sum of a non-cyclic and a cyclic elementary mode.

In what follows, we deal with a reaction scheme describing glycolysis and gluconeogenesis (see Fig. 8). By the algorithm outlined in the previous section, we obtain nine irreversible elementary modes,

$$\begin{aligned} \mathbf{f}^{(1)} &= (1 \ 0 \ 1 \ 1 \ 0 \ 1 \ 1 \ 0 \ 0 \ 1 \ 0)^T, & \mathbf{f}^{(2)} &= (1 \ 0 \ 1 \ 1 \ 0 \ 1 \ 1 \ 1 \ 0 \ 0 \ 1)^T, \\ \mathbf{f}^{(3)} &= (0 \ 1 \ -1 \ 0 \ 1 \ -1 \ 0 \ 1 \ 1 \ -1 \ 0)^T, & \mathbf{f}^{(4)} &= (0 \ 1 \ -1 \ 0 \ 1 \ -1 \ 0 \ 0 \ 1 \ 0 \ -1)^T, \\ \mathbf{f}^{(5)} &= (0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 1 \ 0 \ -1 \ 1)^T, & \mathbf{f}^{(6)} &= (0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 1 \ 0 \ 1 \ 1 \ -1)^T, \\ \mathbf{f}^{(7)} &= (1 \ 1 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0)^T, & \mathbf{f}^{(8)} &= (0 \ 0 \ 0 \ 1 \ 1 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0)^T, \\ \mathbf{f}^{(9)} &= (0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 1 \ 1 \ 1 \ 0 \ 0)^T. \end{aligned} \quad (6.6a-i)$$

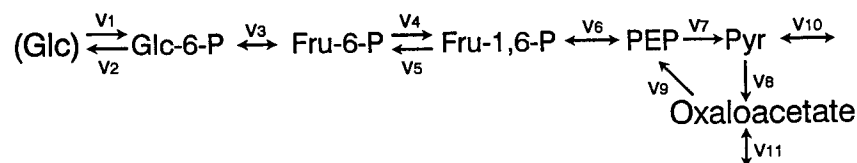


Fig. 8. Simplified reaction scheme of glycolysis and gluconeogenesis. Glucose is treated as external metabolite. Reaction 10 represents both the interconversion of pyruvate and lactate or acetyl-CoA, and of pyruvate and related amino acids.

$\mathbf{f}^{(1)}$ represents the glycolytic pathway. $\mathbf{f}^{(2)}$ stands for the anaplerotic route from glucose to oxaloacetate, which serves to replenish citric acid cycle intermediates. $\mathbf{f}^{(3)}$ and $\mathbf{f}^{(4)}$ correspond to gluconeogenesis pathways emerging from lactate (or amino acids as alanine and serine) and from oxaloacetate (or amino acids such as aspartate and threonine), respectively. $\mathbf{f}^{(5)}$ represents the anaplerotic route from lactate and several amino acids to oxaloacetate. $\mathbf{f}^{(6)}$ can be interpreted as part of the synthesis of several amino acids (such as alanine) under lack of glucose. $\mathbf{f}^{(7)}$, $\mathbf{f}^{(8)}$ and $\mathbf{f}^{(9)}$ are futile cycles (for definition see [5,7,10,13]). The vector $\mathbf{f}^{(7)}$ is no cyclic mode in the sense of Definition 3.4. This is because glucose is here treated as external metabolite.

7. Discussion

The present analysis provides a tool for detecting essential structural features of any given biochemical system not just by inspecting the reaction scheme, but by algebraically analysing the stoichiometry matrix. This method widens the approach of calculating null-space vectors to that matrix. As all elementary modes fulfill the steady-state equation (2.2), they are situated in this null-space, the dimension of which can be less or greater than the number of elementary modes. In the latter case, some of its basis vectors principally cannot represent biochemically meaningful fluxes.

An important application of the null-space matrix \mathbf{K} is, amongst others, Metabolic Control Analysis [6,9], where \mathbf{K} is frequently used to determine control coefficients, based on the generalized summation theorems of Metabolic Control Analysis presented by Reder [16]. Accordingly, computer programs for analysing control properties [14,23] include routines for computing matrix \mathbf{K} . It could be of interest to use elementary modes instead of the columns of \mathbf{K} , to facilitate interpretation of the generalized summation theorems in terms of biochemical functioning.

It is worth noting that contrary to the the null-space matrix \mathbf{K} , the elementary modes are uniquely determined. Since they meet the condition of simplicity and have to reflect the decomposability of the null-space if \mathbf{K} can be transformed into block-diagonal form, we suppose that they relate in some way to the representation (2.5) of \mathbf{K} . This relationship deserves further study in the future.

The definition of cycles as given in the present paper (Definition 3.4) obviously covers futile cycles (also called substrate cycles) [5,7,10,13]. A special case are cycles contained in enzyme cascades [8]. The fact that the definition does not apply to moiety-conserved cycles [17] (e.g., systems preserving the sum of ATP and ADP, or catalytic cycles preserving the sum of enzyme and enzyme-substrate complex), may indicate that usage of the term cycle is somewhat misleading in those situations.

From another point of view, a distinction between cyclic and non-cyclic routes in reaction systems was also made by Feinberg [4], who used the terms circuit and tree. Interestingly, moiety-conserved cycles are no circuits either in that formalism.

It may occur that for a given reaction scheme, some reactions are represented by zero entries in all elementary modes calculated, such as for the pyruvate export reaction in the scheme developed and studied in [12]. If these reactions are known, from experiment, to have non-zero net fluxes, this outcome shows that the biochemical system is not consistently modelled by the scheme. Nevertheless, there do exist reactions that have zero net fluxes independently of the values of kinetic parameters (strictly detailed balanced reactions [22]).

The example of glycolysis and gluconeogenesis shows that the elementary modes relate to the biochemical functions of the network. Therefore, their number may be an important index characterizing biochemical systems. It indicates the richness of the system considered, by showing the variety of its realizable functions. Which of these functions are operative or in what proportions they operate simultaneously is determined by the extent of inhibition and activation of enzymes, i.e., by the actual values of kinetic parameters, which we have not considered in our structural analysis. In consequence of the above reasoning, we believe that on reducing the number of variables in kinetic models (for instance, in applying the quasi-steady state approximation), one should put attention to maintain the number of elementary modes.

Acknowledgements

The authors would like to thank Dr D. A. Fell (Oxford) and two anonymous referees for helpful comments.

References

- [1] Clarke B. L., Stability of complex reaction networks. In *Adv. in Chem. Phys.* **43**, ed. by Prigogine I. and Rice S. A. (Wiley, New York, 1980) pp. 1–216.
- [2] Clarke B. L., Stoichiometric network analysis. *Cell Biophys.* **12** (1988) pp. 237–253.
- [3] Edelstein-Keshet L., *Mathematical Models in Biology* (Random House, New York, 1988).
- [4] Feinberg M., Necessary and sufficient conditions for detailed balancing in mass action systems of arbitrary complexity. *Chem. Eng. Sci.* **44** (1989) pp. 1819–1827.
- [5] Fell D. A., Substrate cycles: theoretical aspects of their role in metabolism. *Comments on Theor. Biol.* **6** (1990) pp. 1–14.
- [6] Fell D. A., Metabolic control analysis: a survey of its theoretical and experimental development, *Biochem. J.* **286** (1992) pp. 313–330.

- [7] Fell D. A., The analysis of flux in substrate cycles. In *Modern Trends in Biothermokinetics*, ed. by Schuster S., Rigoulet M., Ouhabi M. and Mazat J. P. (Plenum Press, New York and London, 1993), pp. 97–101.
- [8] Goldbeter A. and Koshland Jr. D. E., Ultrasensitivity in biochemical systems controlled by covalent modification. *J. Biol. Chem.* **259** (1984) pp. 14441–14447.
- [9] Heinrich R., Rapoport S. M. and Rapoport T. A., Metabolic regulation and mathematical models, *Prog. Biophys. Mol. Biol.* **32** (1977) pp. 1–82.
- [10] Hers H. G. and Hué L., Gluconeogenesis and related aspects of glycolysis. *Ann. Rev. Biochem.* **52** (1983) pp. 617–653.
- [11] Horn F. and Jackson R., General mass action kinetics. *Arch. Rational Mech. Anal.* **47** (1972) pp. 81–116.
- [12] Joshi A. and Palsson B. O., Metabolic dynamics in the human red cell. I. A comprehensive kinetic model. *J. Theor. Biol.* **141** (1989) pp. 515–528.
- [13] Leiser J. and Blum J. J., On the analysis of substrate cycles in large metabolic systems, *Cell Biophys.* **11** (1987) pp. 123–138.
- [14] Letellier T., Reder C. and Mazat J. P., Control: Software for the analysis of the control of metabolic networks. *Comp. Appl. Biosci.* **7** (1991) pp. 383–390.
- [15] Nožička F., Guddat J., Hollatz H. and Bank B., *Theorie der linearen parametrischen Optimierung* (Akademie-Verlag, Berlin, 1974).
- [16] Reder C., Metabolic control theory: A structural approach. *J. Theor. Biol.* **135** (1988) pp. 175–201.
- [17] Reich J. G. and Selkov E. E., *Energy Metabolism of the Cell* (Academic Press, London, 1981).
- [18] Rockafellar R., *Convex Analysis* (Princeton Univ. Press, Princeton, 1970).
- [19] Schuster R. and Schuster S., Refined algorithm and computer program for calculating all non-negative fluxes admissible in steady states of biochemical reaction systems with or without some flux rates fixed. *Comp. Appl. Biosci.* **9** (1993) pp. 79–85.
- [20] Schuster R., Schuster S. and Holzhütter H. G., Simplification of complex kinetic models used for the quantitative analysis of nuclear magnetic resonance or radioactive tracer studies, *J. Chem. Soc. Faraday Trans.* **88** (1992) pp. 2837–2844.
- [21] Schuster S. and Heinrich R., Minimization of intermediate concentrations as a suggested optimality principle for biochemical networks. I. Theoretical analysis. *J. Math. Biol.* **29** (1991) pp. 425–442.
- [22] Schuster S. and Schuster R., Detecting strictly detailed balanced subnetworks in open chemical reaction networks. *J. Math. Chem.* **6** (1991) pp. 17–40.
- [23] Thomas S. and Fell D. A., A computer program for the algebraic determination of control coefficients in metabolic control analysis. *Biochem. J.* **292** (1993) pp. 351–360.