

On the equivalence of direct mechanisms and structurally minimal pathways

Mate Barany · Botond Bertok · Csanad Imreh ·
L. T. Fan · Ferenc Friedler

Received: 24 November 2011 / Accepted: 7 January 2012 / Published online: 31 January 2012
© Springer Science+Business Media, LLC 2012

Abstract A reaction-pathway identification procedure has two distinct phases. The first phase enumerates exhaustively the feasible candidate pathways, and the second phase identifies the ultimate feasible pathway or pathways among them. Probably the most efficient way to execute the first phase is to algorithmically generate the networks of feasible candidate pathways from a predefined set of plausible elementary reactions. The available algorithmic methods for this purpose can be roughly grouped into two major classes, one based on graph theory and the other on linear algebra. Both classes of methods consider any chemical reaction system as a network of elementary reactions, thereby implying that the two classes are interrelated. This paper studies the linear algebraic concept termed direct mechanism introduced in the mid-eighties and the graph-theoretical concept termed structurally minimal pathway introduced two decades later. Herein, it has been formally proven that the two concepts are equivalent.

Keywords Pathway · Mechanism · Linear algebra · Graph theory

M. Barany · B. Bertok (✉) · F. Friedler
Department of Computer Science and Systems Technology, University of Pannonia, Egyetem u. 10,
8200 Veszprém, Hungary
e-mail: bertok@dcs.uni-pannon.hu

F. Friedler
e-mail: friedler@dcs.uni-pannon.hu

C. Imreh
Department of Computer Algorithms and Artificial Intelligence, University of Szeged, Árpád tér 2,
6720 Szeged, Hungary

L. T. Fan
Department of Chemical Engineering, Kansas State University, Manhattan, KS 66506, USA

1 Introduction

Reaction-pathway identification plays a key role in exploring the kinetics of catalytic reactions. For example, it is essential for understanding the effects of external conditions on the rates of catalytic reactions. Such understanding can be a basis for various applications, especially for designing novel industrial chemical processes. Two distinct phases exist in the procedure for identifying or determining a reaction pathway. In the first phase, all the feasible candidate pathways are enumerated, while in the second phase, the ultimate feasible pathway or pathways are identified among them. The most effective way to accomplish the first phase is probably through the algorithmic generation of networks of feasible candidate pathways from a given set of plausible elementary reactions for which several methods have been proposed.

The above mentioned methods can be roughly grouped into two classes; one is based on graph theory (see, e.g., [1,2]) and the other, on linear algebra (see, e.g., [5,8]). While these classes of methods differ in nature, either class considers a chemical reaction system as a network comprising elementary chemical reactions linked to one another through shared reactants, thereby implying that they might be related to each other. The present work proves that the concept of direct mechanisms resulting from the method rooted in linear algebra and that of structurally minimal pathways resulting from the method rooted in graph theory are identical. For clarification, both methods are illustrated with a specific chemical reaction system, namely catalytic ammonia synthesis; see the “Appendix”.

2 Direct mechanisms

Introduced herein is the linear algebraic method developed by Happel and Sellers [5]. A mathematical discourse of the method entails the formal definition of the chemical reaction system under consideration. Happel and Sellers [5] have defined two vector spaces, an S -dimensional space of chemical reaction mechanisms and a Q -dimensional space of chemical reactions. These two vector spaces are related to each other since each mechanism \mathbf{m} gives rise to a unique reaction $R(\mathbf{m})$, wherein R is a function transforming the mechanism into the reaction. R is linear: the reactions in a chemical reaction system are additive, and thus, the reaction associated with combined mechanisms $\mathbf{m}_1 + \mathbf{m}_2$ is $R(\mathbf{m}_1) + R(\mathbf{m}_2)$. The simplest kind of mechanism consists of a one-step molecular interaction and is termed a step. Any mechanism is a combination of such steps. Each of the steps produces one of the elementary reactions forming a basis for the space of all reactions involved in the chemical reaction system.

The formalization presented in [5] is as follows: Denote the species contained in the chemical reaction system by a_1, a_2, \dots, a_A and the elementary reactions among these species by the S vectors as given below:

$$\begin{aligned}
 \mathbf{r}_1 &= \gamma_{11}a_1 + \gamma_{12}a_2 + \cdots + \gamma_{1A}a_A \\
 \mathbf{r}_2 &= \gamma_{21}a_1 + \gamma_{22}a_2 + \cdots + \gamma_{2A}a_A \\
 &\vdots \quad \quad \quad \vdots \\
 \mathbf{r}_S &= \gamma_{S1}a_1 + \gamma_{S2}a_2 + \cdots + \gamma_{SA}a_A
 \end{aligned} \tag{1}$$

where the γ 's are stoichiometric coefficients. Presumably, each elementary reaction has at least one positive coefficient as well as at least one negative coefficient, and every reaction is reversible.

The elementary reactions in Eq. (1) may be linearly dependent. The maximum number of linearly independent reaction vectors in a linearly independent subset is denoted by Q . This subset provides a basis for a Q -dimensional vector space, termed the reaction space. Furthermore, the molecular interaction yielding reaction \mathbf{r}_i is denoted by step s_i . Let mechanism \mathbf{m} be any linear combination of steps in the form

$$\mathbf{m} = \sigma_1 \mathbf{s}_1 + \sigma_2 \mathbf{s}_2 + \cdots + \sigma_S \mathbf{s}_S \quad (2)$$

where the coefficients, σ_i , are real numbers signifying the rate of occurrence of s_i . The S -dimensional vector space comprising the set of all such mechanisms is called the mechanism space. The reaction, \mathbf{r} , corresponding to the mechanism, \mathbf{m} , can be obtained by applying the linear function, R , to the above equation, thereby resulting in

$$\mathbf{r} = \sigma_1 \mathbf{r}_1 + \sigma_2 \mathbf{r}_2 + \cdots + \sigma_S \mathbf{r}_S \quad (3)$$

The equations of Eq. (1) can be substituted into this expression, thereby rendering it possible to express it by the following explicit linear combination:

$$\mathbf{r} = R(\mathbf{m}) = \left(\sum_{i=1}^S \sigma_i \gamma_{i1} \right) a_1 + \left(\sum_{i=1}^S \sigma_i \gamma_{i2} \right) a_2 + \cdots + \left(\sum_{i=1}^S \sigma_i \gamma_{iA} \right) a_A \quad (4)$$

The matrix form of this equation where the stoichiometric coefficients of all the elementary reactions are given by a single matrix is as follows:

$$\mathbf{r} = R(\mathbf{m}) = (\sigma_1, \sigma_2, \dots, \sigma_S)^T \begin{pmatrix} \gamma_{11} & \gamma_{12} & \cdots & \gamma_{1A} \\ \gamma_{21} & \gamma_{22} & \cdots & \gamma_{2A} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{S1} & \cdots & \cdots & \gamma_{SA} \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_A \end{pmatrix} \quad (5)$$

Species in a chemical reaction system can be grouped into two classes. One comprises terminal species including starting reactants and final products. The other comprises intermediates that do not belong to the terminal species. In a steady-state mechanism, the concentrations of all intermediates are presumed to be constant, thus implying that the net rate of production of every intermediate is zero.

The intermediate species are denoted by a_1, a_2, \dots, a_I , and the terminal species, by $a_{I+1}, a_{I+2}, \dots, a_{I+T}$ where $I + T = A$. Consequently, the first I coefficients in the right-hand side of Eq. (4) are zero. Thus, Horiuti [6] has introduced a characterization for such a mechanism as one whose coefficients $\sigma_1, \sigma_2, \dots, \sigma_S$ satisfy the I linear equations expressed as

$$(\sigma_1, \sigma_2, \dots, \sigma_S) \begin{pmatrix} \gamma_{11} & \cdots & \gamma_{1I} \\ \vdots & \ddots & \vdots \\ \gamma_{S1} & \cdots & \gamma_{SI} \end{pmatrix} = (0, 0, \dots, 0) \quad (6)$$

If the rank of the $S \times I$ matrix in the above equation is denoted by H , the dimension of the space of all steady-state mechanisms, P , is equal to $(S - H)$. The values of S , H , and P and the relations among them can be determined by simply treating them as dimensions of vector spaces and resorting to the concepts of basis and linear independence of vectors. In linear algebra, a basis for a vector space is a sequence of vectors that form a set, which is linearly independent and spans the space. Since the dimension of a space is equal to the number of elements in a basis, every steady-state mechanism can uniquely be expressed in terms of P steady-state mechanisms. While this approach uniquely represents each steady-state mechanism, it does not provide a valid classification from a chemical point of view since the choice of basis is arbitrary and is not dictated, in general, by any consideration of chemistry. There exists, however, a unique collection of mechanisms in every chemical reaction system, called direct mechanisms. Let \mathbf{m} be a mechanism and \mathbf{r} be the reaction, which it produces. Mechanism \mathbf{m} is defined as direct if it is minimal in the sense that, if one step is omitted, then no such a mechanism can be formed from any linear combination of the remaining steps leading to \mathbf{r} .

In every chemical reaction system, there exists at least one such set of direct mechanisms providing a basis for the vector space of all mechanisms of the given system. Usually, a proper subset of all direct mechanisms is sufficient to span the space of all potential mechanisms, i.e., there are more direct mechanisms than basis elements for a given system. While this implies that linear dependence relations can exist among the direct mechanisms, they differ chemically. Moreover, a linear algebraic basis for a system is ambiguous while the set of direct mechanisms is a uniquely defined attribute of the system.

If the direct mechanisms are linearly independent in the chemical reaction system under consideration, a method of Happel and Sellers [5] is able to determine each of them. This algorithm and a more detailed description of the method for determining direct mechanisms are found in [5]. As mentioned previously, the method is illustrated with the catalytic synthesis of ammonia, the details of which are given in the “Appendix”.

3 Structurally minimal pathways

An approach entirely different from the one presented in the preceding section is introduced here for enumerating reaction mechanisms. Stoichiometrically exact candidate pathways of a complex reaction can be determined through the synthesis of networks of plausible steps constituting such pathways.

Process synthesis problems are ubiquitous in chemical and allied industries. Process synthesis is of major interest due to its utmost practical importance; nevertheless, it is extremely difficult to execute. Process synthesis gives rise to a complex

combinatorial optimization problem that is generally considered to be highly convoluted and computationally demanding.

Reaction-pathway identification can be defined as a class of process synthesis problems where each species consists of a finite number of chemical elements in a fixed ratio, and these chemical elements are conserved throughout the process. The products of an elementary reaction comprise exactly the same chemical elements as the starting reactants, and the products of an overall reaction contain exactly the same components as the starting reactants.

The reaction-pathway identification is defined by the quadruple (\mathbf{E}, O, M, Q) where $Q = \{q_1, q_2, \dots, q_h\}$ is the finite ordered set of the components of the species; and $M = \{a_1, a_2, \dots, a_l\}$, the finite ordered set of species, each of which is defined by an $\mathbf{a}_j = (a_{1,j}, a_{2,j}, \dots, a_{h,j}) \in (R_0^+)^h$ vector of nonnegative numbers with $a_{k,j}$ denoting the quantity of component q_k in the species \mathbf{a}_j ($k = 1, 2, \dots, h$). The overall reaction denoted by \mathbf{E} is given by the l -dimension vector of real numbers, $\mathbf{E} = (E_1, E_2, \dots, E_l) \in R^l$, where E_j signifies the difference of the production and consumption rates of the species \mathbf{a}_j ($j = 1, 2, \dots, l$) by the overall reaction. Species \mathbf{a}_j is a starting reactant of the overall reaction if and only if $E_j < 0$, and it is a final product of the overall reaction if and only if $E_j > 0$. Finally, $O = \{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ is the finite ordered set of elementary reactions where every reaction \mathbf{e}_i is represented by an l -dimensional vector of real numbers, $\mathbf{e}_i = (e_{1,i}, e_{2,i}, \dots, e_{l,i}) \in R^l$, where $e_{j,i}$ indicates the difference of the rates of consumption and production of the species, \mathbf{a}_j ($j = 1, 2, \dots, l$), by elementary reaction i . Presumably,

$$Q \cap M = M \cap O = O \cap Q = \emptyset \text{ and } \mathbf{E} \notin Q \cup M \cup O \quad (7)$$

The discourse in the preceding paragraph implies that the combinatorial nature of reaction-pathway identification makes it difficult to execute. In constituting a mechanism or network directed from the starting reactants towards the final products or vice versa, each step contributes the forward, reverse or no step to the network. For example, a network comprising only ten steps results in $(3^{10} - 1) = 59048$ possible combinations.

Every feasible process has common combinatorial properties that can be exploited in process synthesis. As such, Friedler et al. [3,4] have proposed a method that is capable of solving realistic and exceedingly complex industrial process synthesis problems. The method has been adapted Fan et al. [1] for reaction-pathway identification.

To exploit the combinatorial properties of any feasible network, it is necessary to represent the structure of the network unambiguously. The elementary reactions performed by the steps are directed, and therefore, every network representing a reaction pathway, including these reactions, can be represented by directed graphs. For this purpose, Fan et al. [1] have proposed to adapt a unique directed bipartite graph, specifically the P-graph, for reaction-pathway identification.

In the graphical representation of a P-graph, elementary reactions are depicted by horizontal bars while chemical and active species by circles. If a chemical species is a reactant to an elementary reaction, the vertex representing the species is linked to the vertex representing the elementary reaction by an arc. Similarly, if a chemical species

is a product from an elementary reaction, then the vertex representing the elementary reaction is linked by an arc to the vertex representing the species.

The set of chemical or active species is denoted by $m \subseteq M$; and the set of elementary reactions under consideration, by $o \subseteq O$, where

$$m = \{a_j \in (R_0^+)^h : j = 1, 2, \dots, l'\} \quad l' \leq l \quad (8)$$

and

$$o = \{\mathbf{e}_i \in R^l : i = 1, 2, \dots, n'\} \quad n' \leq n \quad (9)$$

An (m, o) P-graph representing a reaction network can be defined formally as follows: The set of vertices is $m \cup o$, and the set of arcs is $A_1 \cup A_2$, where

$$A_1 = \{(a_j, \mathbf{e}_i) : a_j \in m, \mathbf{e}_i \in o, \mathbf{e}_{j,i} < 0\} \quad (10)$$

and

$$A_2 = \{(\mathbf{e}_i, a_j) : \mathbf{e}_i \in o, a_j \in m, \mathbf{e}_{j,i} > 0\} \quad (11)$$

An (m', o') P-graph is a subgraph of the (m'', o'') P-graph, i.e., $(m', o') \subseteq (m'', o'')$, if $m' \subseteq m''$ and $o' \subseteq o''$. The union of P-graphs (m', o') and (m'', o'') is defined to be the P-graph $(m' \cup m'', o' \cup o'')$.

The reaction pathway leading from the starting reactants to the final products of the overall reaction of interest is *combinatorially feasible*, if it satisfies the following axioms [1]:

- (T1) Every final product is represented in the network.
- (T2) Every starting reactant is represented in the network.
- (T3) Each reaction step represented in the network is defined a priori.
- (T4) Every active species represented in the network has at least one path leading to the final product of the overall reaction.
- (T5) Every chemical or active species represented in the network must be a reactant for or a product from at least one reaction step represented in the network.
- (T6) A reactant of any elementary reaction represented in the reaction network is a starting reactant, if it is not produced by any reaction step represented in the network.
- (T7) The network includes at most either the forward or the reverse step of each elementary reaction represented in the network.

Based on the above axioms, a set of algorithms has been crafted focusing on the combinatorial properties of feasible reaction networks. Algorithm RPIMSG, which is a slight adaptation of algorithm MSG, generates the maximal structure of the reaction network of interest in polynomial time. The importance of the maximal structure, which is a mathematically rigorous superstructure, has been thoroughly explored and is well-understood [7]. The maximal structure of reaction networks gives rise to algorithm RPISSG, which is the adaptation of algorithm SSG.

Not every combinatorially feasible structure constitutes a *feasible pathway*. A feasible pathway is a pathway satisfying the following six axioms;

- (R1) Every final product is totally produced by the reaction steps represented in the pathway.
- (R2) Every starting reactant is totally consumed by the reaction steps represented in the pathway.
- (R3) Every active intermediate produced by any reaction step represented in the pathway is totally consumed by one or more reaction steps in the pathway, and every active intermediate consumed by any reaction step represented in the pathway is totally produced by one or more reaction steps in the pathway,
- (R4) All reaction steps represented in the pathway are defined a priori.
- (R5) The network representing the pathway is acyclic.
- (R6) At least one elementary-reaction step represented in the pathway activates a starting reactant.

It can be seen from these axioms that feasible pathways must satisfy stronger constraints than combinatorially feasible pathways. This is obvious, since axioms (T1) through (T7) are obtained by relaxing the conditions imposed by axioms (R1) through (R6). Consequently, every feasible pathway is a combinatorially feasible pathway, while the converse is not necessarily true. The complete set of feasible pathways of a given problem can be generated with a branch-and-bound-like algorithm termed PBT.

A P-graph is termed a *structurally minimal pathway* or “*independent pathway*” [1] if it represents a feasible pathway and none of its proper subgraphs can represent a feasible pathway. For a more exact definition of structurally minimal pathways it is necessary to formalize axioms (R1) through (R6) mathematically.

The problem definition implies that a set o of elementary reactions satisfies axioms (R1) through (R3) if and only if there exists a positive coefficient λ_i for each elementary reaction $\mathbf{r}_i \in o$ such that

$$\exists \lambda = (\lambda_1, \lambda_2, \dots, \lambda_n) : \sum_{\mathbf{r}_i \in o} \lambda_i \mathbf{r}_i = \mathbf{E}, \quad \mathbf{r}_i \in o \iff \lambda_i > 0 \quad (12)$$

The above equation implies that the system must be at steady state according to axioms (R1) through (R3). Axiom (R4) can be formulated as

$$o \subseteq O \quad (13)$$

and finally, axiom (R5) can be formulated as

$$\nexists o' : o' \subseteq o, \quad o' \neq \emptyset, \quad \exists \lambda' = (\lambda'_1, \lambda'_2, \dots, \lambda'_n) : \sum_{\mathbf{r}_i \in o'} \mathbf{r}_i \lambda'_i = 0, \quad \mathbf{r}_i \in o' \iff \lambda'_i > 0 \quad (14)$$

From a mathematical point of view, an (m, o) P-graph is termed structurally minimal pathway, if it satisfies the following statement: there exist such positive coefficients

for the reaction steps included in o , that the overall reaction can be written as a linear combination of the elements of o with the given coefficients. Moreover (m, o) is minimal in the sense that it has no proper subgraphs satisfying this criteria. Algorithm PBT is capable of finding every structurally minimal pathway of a given reaction [1].

As mentioned in the introductory section, the method based on P-graphs yielding the stoichiometrically exact candidate feasible pathways is also illustrated with the catalytic ammonia synthesis; see the “Appendix” for details.

4 Proof of the equivalence

This section presents the equivalence of the term, “direct mechanisms”, introduced by Happel and Sellers [5] and the term, “structurally minimal”, i.e. “independent pathways”, introduced by Fan et al. [1]. First, it is shown that the overall reaction produced by a direct mechanism can be expressed by exactly one linear combination of the elementary reactions produced by the steps constituting the direct mechanism.

Theorem 1 *Let \mathbf{m} be a direct mechanism yielding the overall reaction \mathbf{E} . Then, the overall reaction can be expressed by exactly one linear combination of the elementary reaction vectors, $\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_n$, generated by the steps $\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_n$ constituting the direct mechanism, \mathbf{m} .*

Proof Let \mathbf{m} be a direct mechanism yielding overall reaction \mathbf{E} . Let $\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_n$ be the steps constituting the direct mechanism, \mathbf{m} , and let $\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_n$ be elementary reactions generated by the steps, $\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_n$. Then, \mathbf{E} can be expressed as a linear combination of $\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_n$ with coefficients $\lambda_1, \lambda_2, \dots, \lambda_n$:

$$\lambda_1 \mathbf{r}_1 + \lambda_2 \mathbf{r}_2 + \dots + \lambda_n \mathbf{r}_n = \mathbf{E} \quad (15)$$

Suppose that the overall reaction can be written as another linear combination of these vectors with coefficients $\lambda_1^*, \lambda_2^*, \dots, \lambda_n^*$, as

$$\lambda_1^* \mathbf{r}_1 + \lambda_2^* \mathbf{r}_2 + \dots + \lambda_n^* \mathbf{r}_n = \mathbf{E} \quad (16)$$

and that there exists at least one such index k for which λ_k differs from λ_k^* , say

$$\exists k, \quad \lambda_k - \lambda_k^* \neq 0 \quad (17)$$

For $i = 1, 2, \dots, n$, let ε_i denote the difference, $\lambda_i - \lambda_i^*$. Then, the linear combination of the elementary reaction vectors, $\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_n$, with coefficients $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ is the null vector, i.e.,

$$\varepsilon_1 \mathbf{r}_1 + \varepsilon_2 \mathbf{r}_2 + \dots + \varepsilon_n \mathbf{r}_n = \mathbf{E} - \mathbf{E} = 0 \quad (18)$$

This implies that the elementary reaction vectors are not linearly independent. Hence, one of the elementary reactions, $\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_n$, can be expressed as a linear combination of the others. For example, for r_k where λ_k differs from λ_k^* (and thus ε_k is nonzero):

$$\varepsilon_k \mathbf{r}_k = - \sum_{\substack{i=1 \\ i \neq k}}^n \varepsilon_i \mathbf{r}_i \quad (19)$$

The above expression signifies that mechanism \mathbf{m} is not a direct mechanism: The overall reaction \mathbf{E} could be expressed as a linear combination of the reaction vectors $\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_{k-1}, \mathbf{r}_{k+1}, \dots, \mathbf{r}_n$ omitting reaction \mathbf{r}_k as follows:

$$\lambda_1 \mathbf{r}_1 + \lambda_2 \mathbf{r}_2 + \dots - \frac{\lambda_k}{\varepsilon_k} \sum_{\substack{i=1 \\ i \neq k}}^n \varepsilon_i \mathbf{r}_i + \dots + \lambda_n \mathbf{r}_n = \mathbf{E} \quad (20)$$

This contradiction proves the theorem. \square

Corollary 1 *To define a direct mechanism, it is sufficient to define the set of elementary reactions constituting the direct mechanism.*

This corollary implies that it is unnecessary to define the coefficients for the elementary reactions to uniquely define a direct mechanism; it is sufficient to define the set of elementary reactions constituting the direct mechanism. It is significant that Corollary 1 provides the link between the methodology of Happel and Sellers [5] and the methodology of Fan et al. [1]. Combining this result with Theorem 2 presented below will prove that the direct mechanisms and the structurally minimal pathways of a chemical system are identical.

Theorem 2 *There is no proper subset of elementary reactions forming a structurally minimal pathway that can constitute a direct mechanism.*

Proof Suppose that for a vector \mathbf{r} a given set of vectors $\{\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_k, \mathbf{r}_{k+1}, \dots, \mathbf{r}_n\}$ is minimal in the sense that the vector, \mathbf{r} , cannot be written as a linear combination of vectors of any proper subset of $\{\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_k, \mathbf{r}_{k+1}, \dots, \mathbf{r}_n\}$ with positive coefficients, i.e.,

$$\mathbf{r} = \sum_{i=1}^n \lambda_i \mathbf{r}_i, \quad \forall \lambda_i > 0 \quad (21)$$

\mathbf{r} is minimal in the sense that there is no proper subset of $\{\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_k, \mathbf{r}_{k+1}, \dots, \mathbf{r}_n\}$ satisfying Eq. (14). Now suppose that by relaxing the positivity constraint and enabling negative coefficients, \mathbf{r} can be written as a linear combination of the vectors in $\{\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_k\}$, namely

$$\mathbf{r} = \sum_{i=1}^k \lambda_i^* \mathbf{r}_i \quad \text{where } k < n \quad (22)$$

Let δ be

$$\delta = \min_{\lambda_i^* < 0} \left\{ \frac{\lambda_i}{|\lambda_i^*|} \right\} \quad (23)$$

This δ is positive: since if $\lambda_i \geq 0$, then

$$\frac{\lambda_i}{|\lambda_i^*|} \geq 0 \quad (24)$$

Multiplying both sides of Eq. (22) by δ gives

$$\delta \mathbf{r} = \sum_{i=1}^k \delta \lambda_i^* \mathbf{r}_i \quad (25)$$

The sum of Eqs. (21) and (25) yields

$$\mathbf{r}(1 + \delta) = \sum_{i=1}^k (\lambda_i + \delta \lambda_i^*) \mathbf{r}_i + \sum_{i=k+1}^n \lambda_i \mathbf{r}_i \quad (26)$$

In light of the definition of δ , in the first sum, where the minimum is attained, one of the coefficients $(\lambda_i + \delta \lambda_i^*)$ will be zero. By denoting the index of this coefficient by j , we have

$$\delta = \frac{\lambda_j}{|\lambda_j^*|} \quad \text{and} \quad \lambda_j + \frac{\lambda_j}{|\lambda_j^*|} \lambda_j^* = 0 \quad (27)$$

The other coefficients will be nonnegative: it is trivial that

$$\lambda_i + \frac{\lambda_j}{|\lambda_j^*|} \lambda_i^* \geq 0 \quad \text{where} \quad \lambda_i^* \geq 0 \quad (28)$$

and

$$\lambda_i + \frac{\lambda_j}{|\lambda_j^*|} \lambda_i^* = 0 \quad \text{where} \quad \lambda_i^* < 0 \quad (29)$$

Since δ is defined where the minimum is attained, clearly

$$0 < \delta = \frac{\lambda_j}{|\lambda_j^*|} \leq \frac{\lambda_i}{|\lambda_i^*|} \quad \forall i, i \neq j \quad (30)$$

thereby yielding

$$\lambda_i + \frac{\lambda_j}{|\lambda_j^*|} \lambda_i^* \geq 0 \quad \forall i, i \neq j \quad (31)$$

Thus, the vector, \mathbf{r} , can be written as a linear combination of the vectors of this proper subset with positive coefficients; moreover, by dividing both sides with $(1 + \delta)$, a

proper subset of the set, $\{\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_k, \mathbf{r}_{k+1}, \dots, \mathbf{r}_n\}$, is obtained. They are given below.

$$\mathbf{r}(1 + \delta) = \sum_{\substack{i=1 \\ i \neq j}}^k (\lambda_i + \delta \lambda_i^*) \mathbf{r}_i + \sum_{i=k+1}^n \lambda_i \mathbf{r}_i \quad (32)$$

$$\mathbf{r} = \frac{\sum_{\substack{i=1 \\ i \neq j}}^k (\lambda + \delta \lambda_i^*) \mathbf{r}_i + \sum_{i=k+1}^n \lambda_i \mathbf{r}_i}{1 + \delta} \quad (33)$$

This contradicts that $\{\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_n\}$ is structurally minimal and proves the following theorem. \square

Theorem 3 *Given a chemical system at steady state, the set of direct mechanisms and the set of structurally minimal pathways of the system are equivalent.*

Proof This theorem follows directly from Corollary 1 and Theorem 2. \square

5 Concluding remarks

The equivalence of direct mechanisms and structurally minimal pathways have been formally defined. Furthermore, it is mathematically proven.

Acknowledgments Authors acknowledge the support of the Hungarian Research Fund under project OTKA 81493K.

A Appendix

This appendix illustrates the methodologies referred in the text of the article, namely those by Happel and Sellers [5] and Fan et al. [1]. The exhaustively studied ammonia synthesis reaction serves the example for the illustration. The overall reaction produces ammonia from hydrogen and nitrogen, i.e.,



Table 1 lists the set of eleven plausible elementary reactions as defined by Happel and Sellers [5] and with more detailed steps by Fan et al. [2].

Tables 2 and 3 list the species and steps involved, respectively, as denoted by Happel and Sellers [5]. What follows represents the stoichiometric coefficients listed in Table 3 by matrix γ .

Table 1 List of candidate elementary reactions

(1)	$H_2 + \ell \rightleftharpoons H_2\ell$
(2)	$H_2\ell + \ell \rightleftharpoons H\ell + H\ell$
(3)	$N_2 + \ell \rightleftharpoons N_2\ell$
(4)	$N_2\ell + \ell \rightleftharpoons N\ell + N\ell$
(5)	$N_2\ell + H_2\ell \rightleftharpoons N_2H_2\ell + \ell$
(6)	$N_2H_2\ell + \ell \rightleftharpoons NH\ell + NH\ell$
(7)	$N\ell + H\ell \rightleftharpoons NH\ell + N\ell$
(8)	$NH\ell + H\ell \rightleftharpoons NH_2\ell + \ell$
(9)	$NH\ell + H_2\ell \rightleftharpoons NH_3\ell + \ell$
(10)	$NH_2\ell + H\ell \rightleftharpoons NH_3\ell + \ell$
(11)	$NH_3\ell \rightleftharpoons NH_3 + \ell$

Table 2 List of the identifiers of the species

Species	H_2	ℓ	$H_2\ell$	$H\ell$	N_2	$N_2\ell$	$N\ell$	$N_2H_2\ell$	$NH\ell$	$NH_2\ell$	$NH_3\ell$	NH_3
Notation	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}	a_{11}	a_{12}

Table 3 Table of stoichiometric coefficients γ_{ij} 's defining the steps

	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}	a_{11}	a_{12}
s_1	-1	-1	1									
s_2			-1	2								
s_3		-1			-1	1						
s_4		-1				-1	2					
s_5		1	-1			-1		1				
s_6		-1						-1	2			
s_7		1		-1			-1		1			
s_8		1		-1					-1	1		
s_9		1	-1						-1		1	
s_{10}		1		-1						-1	1	
s_{11}		1									-1	1

$$\gamma = \begin{pmatrix} -1 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & -1 & 2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & -1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 & 0 & -1 & 2 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 & 0 & -1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & 1 \end{pmatrix} \quad (\text{A2})$$

Table 4 List of direct mechanisms given by steps and corresponding γ coefficients

	s_1	s_2	s_3	s_4	s_5	s_6	s_7	s_8	s_9	s_{10}	s_{11}
m_1	3	1	1	1			2		2		2
m_2	3	3	1	1			2	2		2	2
m_3	3		1	1			2	1	3	1	3
m_4	3		1		1	1			2		2
m_5	3	2	1		1	1		2		2	2
m_6	3		1	2	3	3	4	2		2	2

Table 4 enumerates the direct mechanisms of the system given by the lists of steps involved and the corresponding coefficients. The matrix of these coefficients is denoted by σ as follows:

$$\sigma = \begin{pmatrix} 3 & 1 & 1 & 1 & 0 & 0 & 2 & 0 & 2 & 0 & 2 \\ 3 & 3 & 1 & 1 & 0 & 0 & 2 & 2 & 0 & 2 & 2 \\ 3 & 0 & 1 & 1 & 0 & 0 & 2 & 1 & 3 & 1 & 3 \\ 3 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 2 & 0 & 2 \\ 3 & 2 & 1 & 1 & 1 & 1 & 0 & 2 & 0 & 2 & 2 \\ 3 & 0 & 1 & 2 & 3 & 3 & 4 & 2 & 0 & 2 & 2 \end{pmatrix} \tag{A3}$$

By resorting to the methodology of Fan et al. [1], two chemical elements and an active site are identified in the system. Set Q given below contains these two elements and the notation of the active site ℓ ;

$$Q = \{q_1, q_2, q_3\} = \{\dot{N}, \dot{H}, \dot{\ell}\} \tag{A4}$$

Set M given below includes the twelve species;

$$M = \{a_1, a_2, \dots, a_{12}\} \tag{A5}$$

In an equivalent yet more detailed formulation, the species are represented in vectorial form where the components of the vectors denote the quantities of the chemical elements in the species; as expressed below.

$$M = \{[0, 2, 0]^T, [0, 0, 1]^T, \dots, [1, 3, 0]^T\} \tag{A6}$$

Finally, the set of species can be given by simply listing the species contained in the chemical reaction system by omitting the notations a_1, a_2, \dots, a_{12} , as follows:

$$M = \{H_2, \ell, \dots, NH_3\} \tag{A7}$$

The following equation shows how the set of elementary reactions are given in the terminology introduced by Fan et al. [1]. Note that unlike in the case of steps, it is necessary to define the elementary reactions in both directions;

$$O = \{e_{1\rightarrow}, e_{2\rightarrow}, \dots, e_{11\rightarrow}, e_{1\leftarrow}, e_{2\leftarrow}, \dots, e_{11\leftarrow}\} \quad (\text{A8})$$

A detailed form of this expression is given below where each elementary reaction step is expressed by the vector of its stoichiometric coefficients.

$$\begin{aligned} O &= \{e_{1\rightarrow} = [-1, -1, 1, 0, 0, 0, 0, 0, 0, 0, 0]^T, \\ e_{2\rightarrow} &= [0, 0, -1, 2, 0, 0, 0, 0, 0, 0, 0]^T, \\ &\dots \\ e_{11\rightarrow} &= [0, 1, 0, 0, 0, 0, 0, 0, 0, -1, 1]^T, \\ e_{1\leftarrow} &= [1, 1, -1, 0, 0, 0, 0, 0, 0, 0, 0]^T, \\ e_{2\leftarrow} &= [0, 0, 1, -2, 0, 0, 0, 0, 0, 0, 0]^T, \\ &\dots \\ e_{11\leftarrow} &= [0, -1, 0, 0, 0, 0, 0, 0, 0, 1, -1]^T \} \end{aligned} \quad (\text{A9})$$

Similarly, the following equation shows the vector representing the overall reaction;

$$\mathbf{E} = [-3, 0, 0, 0, -2, 0, 0, 0, 0, 0, 0, 2]^T \quad (\text{A10})$$

Finally listed below is the set V of structurally minimal pathways (m_1, o_1) , (m_2, o_2) , \dots , (m_6, o_6) .

$$\begin{aligned} V &= \{(m_1 = N_2, H_2, NH_3, \ell, H_2\ell, H\ell, N_2\ell, N\ell, NH\ell, NH_3\ell, \\ o_1 &= \{e_{1\rightarrow}, e_{2\rightarrow}, e_{3\rightarrow}, e_{4\rightarrow}, e_{7\rightarrow}, e_{9\rightarrow}, e_{11\rightarrow}\}), \\ (m_2 &= N_2, H_2, NH_3, \ell, H_2\ell, H\ell, N_2\ell, N\ell, NH\ell, NH_2\ell, NH_3\ell, \\ o_2 &= \{e_{1\rightarrow}, e_{2\rightarrow}, e_{3\rightarrow}, e_{4\rightarrow}, e_{7\rightarrow}, e_{8\rightarrow}e_{10\rightarrow}, e_{11\rightarrow}\}), \\ (m_3 &= N_2, H_2, NH_3, \ell, H_2\ell, N_2\ell, N_2H_2\ell, NH\ell, NH_3\ell, \\ o_3 &= \{e_{1\rightarrow}, e_{3\rightarrow}, e_{4\rightarrow}, e_{7\rightarrow}, e_{8\rightarrow}, e_{9\rightarrow}, e_{10\rightarrow}, e_{11\rightarrow}\}), \\ (m_4 &= N_2, H_2, NH_3, \ell, H_2\ell, N_2\ell, N_2H_2\ell, NH\ell, NH_3\ell, \\ o_4 &= \{e_{1\rightarrow}, e_{3\rightarrow}, e_{5\rightarrow}, e_{6\rightarrow}, e_{9\rightarrow}, e_{11\rightarrow}\}), \\ (m_5 &= N_2, H_2, NH_3, \ell, H_2\ell, H\ell, N_2\ell, N_2H_2\ell, NH\ell, NH_2\ell, NH_3\ell, \\ o_5 &= \{e_{1\rightarrow}, e_{2\rightarrow}, e_{3\rightarrow}, e_{5\rightarrow}, e_{6\rightarrow}, e_{8\rightarrow}, e_{10\rightarrow}e_{11\rightarrow}\}), \\ (m_6 &= N_2, H_2, NH_3, \ell, H_2\ell, H\ell, N_2\ell, N\ell, N_2H_2\ell, NH\ell, NH_2\ell, NH_3\ell, \\ o_6 &= \{e_{1\rightarrow}, e_{3\rightarrow}, e_{4\rightarrow}, e_{5\rightarrow}, e_{6\rightarrow}, e_{7\rightarrow}, e_{8\rightarrow}, e_{10\rightarrow}, e_{11\rightarrow}\}) \} \end{aligned} \quad (\text{A11})$$

Note that as expected, the number of structurally minimal pathways is the same as the number of direct mechanisms.

References

1. L.T. Fan, B. Bertok, F. Friedler, *Comput. Chem.* **26**, 265 (2002)
2. L.T. Fan, B. Bertok, F. Friedler, S. Shafie, Hung, *J. Ind. Chem.* **29**, 71 (2001)
3. F. Friedler, K. Tarjan, Y.W. Huang, L.T. Fan, *Comput. Chem. Eng.* **16**, S313 (1992)
4. F. Friedler, K. Tarjan, Y.W. Huang, L.T. Fan, *Chem. Eng. Sci.* **47**, 1973 (1992)

5. J. Happel, P.H. Sellers, *Adv. Catal.* **32**, 273 (1983)
6. J. Horiuti, *Ann. N. Y. Acad. Sci.* **213**, 5 (1973)
7. Z. Kovacs, Z. Ercsey, F. Friedler, L.T. Fan, *Comput. Chem. Eng.* **24**, 1881 (2000)
8. I. Szalkai, *J. Math. Chem.* **28**, 1 (2000)