

Counting minimal reactions with specific conditions in \mathbb{R}^4

Balázs Szalkai · István Szalkai

Received: 16 May 2010 / Accepted: 23 December 2010 / Published online: 9 January 2011
© Springer Science+Business Media, LLC 2011

Abstract We give the sharp lower bound on the number of minimal reactions when no “parallel” species (isomers or multiples) are allowed and all the species are built up from at most *four* kinds of atoms in Theorem 16. This continues the investigations in Kumar and Pethő (Intern Chem Eng 25:767–769, 1985) through Szalkai and Laflamme (Electr J Comb 5(1), 1998) which results we briefly summarize in the first Section.

Keywords Stoichiometry · Minimal reactions · Linear algebra · Simplexes

1 Introduction and preliminary results

One important application of *linear algebraic simplexes* can be found in chemistry (see e.g. in [1]). Though the present paper reveals some new pure mathematical properties of simplexes, these results can be translated also to the language of chemistry, too.

Let us emphasise in advance that there are at least three different notions of simplexes in use: (*linear-*) *algebraic*, *affine* and *geometric* ones. The first two are defined in Definitions 1 and 10 below while the third one is discussed with chemical applications e.g. in [2] and [3].

Definition 1 (i) $S \subset \mathbb{R}^n$ is (*algebraic*) *simplex* iff S is minimal linear dependent, i.e. S is lin. dependent but each proper subset of it is independent. (ii) For any $\mathcal{H} \subset \mathbb{R}^n$ we

B. Szalkai
Eötvös L. University, Budapest, Hungary
e-mail: bszalkai0@gmail.com

I. Szalkai (✉)
Department of Mathematics, University of Pannonia, POB. 158, Veszprém 8201, Hungary
e-mail: szalkai@almos.uni-pannon.hu

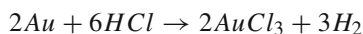
denote by $\mathbf{simp}(\mathcal{H})$ the number of simplexes $S \subset \mathcal{H}$ contained in \mathcal{H} . For $2 \leq k \leq n+1$ we denote by $\mathbf{simp}_k(\mathcal{H})$ the number of k -element simplexes in \mathcal{H} . \square

Simplexes are widely used e.g. in stoichiometry when finding minimal reactions and mechanisms or for finding dimensionless groups in dimensional analysis (see e.g. [4–11]).

Consider for example the notion of minimal reactions. Let the chemical species A_1, A_2, \dots, A_n consist of elements E_1, E_2, \dots, E_m as $A_j = \sum_{i=1}^m a_{i,j} E_i$, ($a_{i,j} \in \mathbb{N}$) for $j = 1, 2, \dots, n$. Writing \underline{A}_j for the vector $[a_{1,j}, a_{2,j}, \dots, a_{m,j}]^T$, we know that there (might) exist a chemical reaction between the species $\{A_j : j \in S\}$ for any $S \subseteq \{1, 2, \dots, n\}$ if and only if the homogeneous linear equation

$$\sum_{j \in S} x_j \underline{A}_j = \underline{0} \quad (1)$$

has nontrivial solution for some $x_j \in \mathbb{R}$, $j \in S$, that is if the vector set $\{\underline{A}_j : j \in S\}$ is *linearly dependent*. Further, the reaction is called **minimal** if for no $T \subsetneq S$ might there be any reaction among the species $\{A_j : j \in T\}$; that is if the vector set $\{\underline{A}_j : j \in T\}$ is *linearly independent* for any $T \subsetneq S$. Of course the reactions obtained in the above way are only possibilities, e.g. the reaction



does not occur under normal conditions.

As a specific example, the species $A_1 = C$, $A_2 = O$, $A_3 = CO$ and $A_4 = CO_2$ determine the vectors $\underline{A}_1 = [1, 0]$, $\underline{A}_2 = [0, 1]$, $\underline{A}_3 = [1, 1]$ and $\underline{A}_4 = [1, 2]$, using the “base” $\{C, O\}$ in \mathbb{R}^2 . The vector set $H = \{\underline{A}_1, \underline{A}_2, \underline{A}_3, \underline{A}_4\}$ contains the simplexes

$$\{\underline{A}_1, \underline{A}_2, \underline{A}_3\}, \quad \{\underline{A}_1, \underline{A}_2, \underline{A}_4\}, \quad \{\underline{A}_1, \underline{A}_3, \underline{A}_4\} \text{ and } \{\underline{A}_2, \underline{A}_3, \underline{A}_4\}.$$

After solving the corresponding Eq. (1), we have the following (complete) list of minimal reactions: $C + O = CO$, $C + 2O = CO_2$, $O + CO = CO_2$ and $C + CO_2 = 2CO$.

For further details and examples see e.g. [10] or [12]. This latter reference further provides a computer algorithm for the pure algebraic problem of finding all the simplexes in a given set of vectors in \mathbb{R}^n , *without repetition*. In [12] the author presents a theoretical approach of the problem [13] also offers an algorithm for this problem.

The main question of the present work is:

Problem 2 For given natural numbers $m, n \in \mathbb{N}$ what is the **minimal** and maximal possible value of $\mathbf{simp}(\mathcal{H})$ and what are the extremal cases assuming that $\mathcal{H} \subset \mathbb{R}^n$, \mathcal{H} spans \mathbb{R}^n and $|\mathcal{H}| = m$? \square

(Clearly the notion and problem above can be stated in matroids, too, see [14].)

In the words of chemistry: *we want lower and upper bounds for the number of minimal reactions whenever both the number of possible atoms and the number of species are given.*

Our main result in the present paper is (see Theorem 16):

Theorem Assuming $\mathcal{H} \subset \mathbb{R}^3$, $|\mathcal{H}| = m \geq 24$ and \mathcal{H} spans \mathbb{R}^3 (i.e. \mathcal{H} is not coplanar) we have

$$\text{simp}(\mathcal{H}) \geq \binom{\lfloor m/2 \rfloor}{3} + \binom{\lceil m/2 \rceil}{3} \tag{2}$$

and equality holds just in the case if \mathcal{H} is contained in two skew (detour) lines and these lines contain $\lfloor m/2 \rfloor$ and $\lceil m/2 \rceil$ many points of \mathcal{H} .

In the language of stoichiometry the above result says:

Theorem There are *at least* $\binom{\lfloor m/2 \rfloor}{3} + \binom{\lceil m/2 \rceil}{3} \approx m^3/24$ many minimal reactions among m species if the species are built up from 4 kinds of atoms and no parallel (multiple) species are allowed.

The minimal number of minimal reactions is achieved if and only if there are four species s_1, s_2, s_3, s_4 with no reaction among them and one half of the other species are linear combinations of s_1, s_2 while the others are of s_3, s_4 . □

The above Theorem continues several previous results which we briefly list below.

The maximal case for $\text{simp}(\mathcal{H})$ in general was solved in [1] and for matroids in [14]:

$$\text{simp}(\mathcal{H}) \leq \binom{m}{n+1}$$

and equality holds if and only if each n -element subsets of \mathcal{H} are independent.¹ (This result is *not* a pure consequence of Sperner’s theorem.)

The minimal case was solved in [1]:

$$b \binom{a+1}{2} + (n-b) \binom{a}{2} \leq \text{simp}(\mathcal{H})$$

assuming $m = an + b$ ($0 \leq b < n$) and \mathcal{H} spans \mathbb{R}^n . Further, equality holds if and only if \mathcal{H} consists of n collections of parallel vectors of sizes a or $a + 1$ (differing by at most one from each other).

For $m \leq 2n$ there are also some other minimal configurations, see [15].

The next step is to *exclude parallel* vectors (i.e. multiple quantities or isomers of species in chemistry). Our first result for the case $n = 3$ appeared in [16]:

Theorem 3 ([16]) For any $\mathcal{H} \subseteq \mathbb{R}^3$ of fixed size not equal to 3, 4, 6 or 7 such that \mathcal{H} spans \mathbb{R}^3 and contains **no** parallel vectors, $\text{simp}(\mathcal{H})$ is minimal if and only if \mathcal{H} is contained in two intersecting planes, one of which is of size 3; i.e. precisely when \mathcal{H}

¹ E.g. the set of vectors $[1, q_i, q_i^2, \dots, q_i^{n-1}]$ where q_1, \dots, q_m are any distinct real numbers—using the result on Vandermonde determinants.

contains 3 linearly independent vectors $\{u_1, u_2, u_3\}$, another vector v coplanar with u_1 and u_2 and the rest $\mathcal{H} \setminus \{u_1, u_2, u_3, v\}$ coplanar with u_2 and u_3 . This implies

$$\binom{m-2}{3} + 1 + \binom{m-3}{2} \leq \text{simp}(\mathcal{H})$$

or, approximatively $\approx \frac{m^3}{6} \leq \text{simp}(\mathcal{H})$. \square

We say many thanks to P. Sellers [17] for pointing out some errors in our above paper, e.g. concerning the case $|\mathcal{H}| = 6$.

In [16] also a general conjecture was settled:

Conjecture 4 (Laflamme-Szalkai, [16]): For the general problem in \mathbb{R}^n regarding the minimum size of $\text{simp}(\mathcal{H})$ where \mathcal{H} is of fixed size, spans \mathbb{R}^n and contains no colinear vectors the minimum is attained **precisely** for the following configurations:

- (i) If n is even, \mathcal{H} contains n linearly independent vectors $\{u_i : i = 1, \dots, n\}$ and the remaining divided as evenly as possible between the planes $\{[u_i, u_{i+1}] : i = 1, 3, \dots, n-1\}$.
- (ii) If n is odd, \mathcal{H} again contains n linearly independent vectors $\{u_i : i = 1, \dots, n\}$, one extra vector in the plane $[u_{n-1}, u_n]$ and finally the remaining vectors divided as evenly as possible between the planes $\{[u_i, u_{i+1}] : i = 1, 3, \dots, n-2\}$ with lower indices having precedence.

In the language of chemistry we say:

Conjecture 5 If the species are built up from n kinds of atoms and no parallel (multiple) species are allowed, then the minimal number of minimal reactions is achieved if and only if

- (i) If n is even, there are n species s_1, s_2, \dots, s_n with no reaction among them (at all), all the n kinds of atoms occur in s_1, s_2, \dots, s_n , and all the remaining species are divided as evenly as possible between the planes $[s_1, s_2], [s_3, s_4], \dots, [s_{n-1}, s_n]$ —in other words: the remaining species are linear combinations $\alpha s_i + \beta s_{i+1}$ of these pairs $\{s_i, s_{i+1}\}$.

In this case the minimal number of minimal reactions is

$$b \binom{\lceil 2m/n \rceil}{3} + (n-b) \binom{\lfloor 2m/n \rfloor}{3} \approx \frac{2m^3}{3n^2} \quad (3)$$

where m denotes the number of species, $m \geq \frac{3}{2}n$ and $m - n = a\frac{n}{2} + b$ ($0 \leq b < \frac{n}{2}$).

- (ii) If n is odd, there are n species s_1, s_2, \dots, s_n with no reaction among them (at all), all the n kinds of atoms occur in s_1, s_2, \dots, s_n , one extra species in the form $\mu s_{n-1} + \nu s_n$, and finally the remaining species are divided as evenly as possible between the planes $[s_1, s_2], [s_3, s_4], \dots, [s_{n-2}, s_{n-1}]$ with lower indices having precedence.

In this case the minimal number of minimal reactions is

$$b \left(\left\lceil \frac{2(m-2)}{n-1} \right\rceil \right) + (n-b) \binom{2(m-2)}{3} + 1 \approx \frac{2m^3}{3n^2} \tag{4}$$

where m denotes the number of species, $m \geq \frac{3}{2}(n-1) + 2$ and $m - (n+1) = a\frac{n-1}{2} + b$ ($0 \leq b < \frac{n-1}{2}$).

The approximations in (3) and (4) are valid for large m . □

In Theorem 16 we prove (i) for $n = 4$, after reducing the problem to \mathbb{R}^3 in the next Section (see the construction in Lemma 12). For dimensions (number of types of atoms) $n \geq 5$ the above Conjecture has not been proved yet.

The remainder part of the present paper is a mathematical justification of Theorem 16—the case $n = 4$ of Conjecture above.

We conclude with some general properties of simplexes:

Claim 6 (o) $\emptyset \notin \mathcal{H}$ can be assumed,

- (i) $S = \{\underline{x}_1, \dots, \underline{x}_k\} \subset \mathbb{R}^n$ is a simplex if and only if for any $i \leq k$ we have: $S \setminus \{\underline{x}_i\}$ is independent and \underline{x}_i is a unique linear combination of the elements of $S \setminus \{\underline{x}_i\}$ where none of the coefficients is 0,
- (ii) multiplying each element of \mathcal{H} by (possible different) scalars $\text{simp}(\mathcal{H})$ does **not** change. □

2 Reducing the dimension

In this Section we reduce the problems $\mathcal{H} \subset \mathbb{R}^n$ from any dimension n to another one $\mathcal{H}' \subset \mathbb{R}^{n-1}$ using the following Remark:

Remark 7 In what follows we assume that $\mathcal{H} \subset \mathbb{R}^n$ does **not** contain parallel vectors (i.e. \mathcal{H} does not contain 2 -element simplexes) and we are interested in the **minimal** value of $\text{simp}(\mathcal{H})$ □

Now we can replace each vector from \mathcal{H} with its *direction* using Claim 6 (ii) (i.e. we map \mathcal{H} into the projective space $P(\mathbb{R}^n)$):

Notation 8 For any $\underline{h} \in \mathbb{R}^n$ we let

$$\Lambda \underline{h} := \{\lambda \cdot \underline{h} : \lambda \in \mathbb{R}, \lambda \neq 0\}$$

and for $\mathcal{H} \subset \mathbb{R}^n$

$$\Lambda \mathcal{H} := \{\Lambda \underline{h} : \underline{h} \in \mathcal{H}\}. \tag{5}$$

Let us now investigate the intersection

$$\Lambda \mathcal{H} \cap \mathcal{P} \tag{5}$$

for a suitable $n - 1$ -dimensional affine hyperplane $\mathcal{P} \subset \mathbb{R}^n$ which is *not* parallel to any $\underline{h} \in \mathcal{H}$. By the definition of $\Lambda\mathcal{H}$, there is a bijective correspondence between \mathcal{H} and the intersection above in (5), which is an $n - 1$ -dimensional subspace of \mathbb{R}^n . How can we reformulate the notion of *simplexes* in the intersection (5)?

Cases $n = 3$ and $n = 4$ can be easily seen (for subsets $\mathcal{S} \subseteq \Lambda\mathcal{H} \cap \mathcal{P}$):

Definition 9

- (i) A set of points $\mathcal{S} \subset \mathbb{R}^2$ is an *affine*
 - ▷ 3-element *simplex* iff \mathcal{S} is three colinear points,
 - ▷ 4-element simplex iff \mathcal{S} is any four points but none three of them are colinear,
 - ▷ there are no other affine simplexes in \mathbb{R}^2 .
- (ii) A set of points $\mathcal{S} \subset \mathbb{R}^3$ is an *affine*
 - ▷ 3-element *simplex* iff \mathcal{S} is three colinear points,
 - ▷ 4-element simplex iff \mathcal{S} is any four coplanar points but none three of them are colinear,
 - ▷ 5-element simplex iff \mathcal{S} is any five points but none four are coplanar (and thus none three of them are colinear),
 - ▷ there are no other affine simplexes in \mathbb{R}^3 . □

(This concept could also be defined in projective spaces as we practically projectivized \mathcal{H} above.)

Now, the general definition goes as follows:

Definition 10 Any finite set $\mathcal{S} = \{\underline{s}_1, \underline{s}_2, \dots, \underline{s}_k\} \subset \mathbb{R}^n$ is an *affine simplex* if and only if $k \geq 3$, the set

$$\{\underline{s}_1 - \underline{s}_k, \underline{s}_2 - \underline{s}_k, \dots, \underline{s}_{k-1} - \underline{s}_k\}.$$

is linearly dependent and *no* proper subset of \mathcal{S} is an affine simplex. □

We also have to correspond the elements of \mathcal{H} and $\Lambda\mathcal{H} \cap \mathcal{P}$ in (5):

Definition 11 For fixed \mathcal{P} and any vector \underline{h} we denote

$$\wp(\underline{h}) := \Lambda\underline{h} \cap \mathcal{P}$$

the “image” (or “trace”) of \underline{h} in \mathcal{P} . □

Now we can reduce the dimension of our original problem with the help of the following Lemma:

Lemma 12 For any finite (or countable) set of vectors $\mathcal{H} \subset \mathbb{R}^n$ there is an $n - 1$ -dimensional hyperplane $\mathcal{P} \subset \mathbb{R}^n$ for which, putting

$$\mathcal{H}' := \Lambda\mathcal{H} \cap \mathcal{P}$$

we have $|\mathcal{H}'| = |\mathcal{H}|$ and

$$\mathcal{S} \subset \mathcal{H} \text{ is an algebraic simplex} \iff \wp(\mathcal{S}) \subset \mathcal{H}' \text{ is an affine simplex.}$$

Proof The $n - 1$ dimensional hyperplanes, excluding the $\underline{0}$, have equation $\underline{r} \cdot \underline{x} = c$ for some $c \neq 0$ and $\underline{r} \neq \underline{0}$. Since \mathcal{H} is at most countable we can chose an $\underline{r} \in \mathbb{R}^n$ not orthogonal to any element of \mathcal{H} (since the functions $f_h(\underline{r}) := \underline{r} \cdot \underline{h}$ are continuous). Using the definition of $\Lambda\mathcal{H}$ and the other definitions above, we conclude that \mathcal{P} and \mathcal{H}' work. \square

In what follows we work in \mathbb{R}^3 and write \mathcal{H} instead of \mathcal{H}' for simplicity, further we omit the attribute affine everywhere.

We will use the following version of Theorem 3 using the above (affine) terms:

Theorem 13 ([16]) *For any set of points $\mathcal{H} \subseteq \mathbb{R}^2$ of fixed size not equal to 3, 4, 6 or 7 such that \mathcal{H} is not contained in a single line, $\text{simp}(\mathcal{H})$ is minimal if and only if \mathcal{H} is contained in two intersecting lines:*

$$\mathcal{H} = \{P_1, P_2, \dots, P_m\} \subseteq \ell_1 \cup \ell_2$$

where $\ell_1 = \{P_1, P_2, P_3\}$ and $\ell_2 = \{P_3, P_4, \dots, P_m\}$. \square

By Lemma 12 and Definition 9 the above Theorem is an equivalent form of Theorem 3.

Similarly, for $n = 4$ we have the following equivalent version of Conjecture 4:

Conjecture 14 *For any set of points $\mathcal{H} \subseteq \mathbb{R}^3$ of fixed size $m \geq m_0$ such that \mathcal{H} is not contained in a single plane, $\text{simp}(\mathcal{H})$ is minimal if and only if \mathcal{H} is contained in two skew lines, these lines contain $\lfloor m/2 \rfloor$ and $\lceil m/2 \rceil$ points, respectively.*

This Conjecture will be proved in Theorem 16.

In what follows, we always talk about $\mathcal{H}' \subset \mathbb{R}^3$ instead of $\mathcal{H} \subset \mathbb{R}^4$ and write simply $\mathcal{H} \subset \mathbb{R}^3$.

3 The proof

So $\mathcal{H} \subset \mathbb{R}^4$ i.e. $\mathcal{H}' \subset \mathbb{R}^3$, $|\mathcal{H}| = |\mathcal{H}'| = m$ and

$$\mathcal{H}' \subset \mathbb{R}^3 \text{ is not coplanar} \tag{6}$$

and we write \mathcal{H} instead of \mathcal{H}' from now on.

Definition 15 (i) For any points $A, B, C \in \mathbb{R}^3$ we denote by $[A, B]$ and by $[A, B, C]$ the line and the plane spanned by the points A, B or A, B, C , respectively. Further, we use the notation $[A, B, C, \dots, N]$ for emphasizing the *coplanarity* of the points $A, B, C, \dots, N \in \mathbb{R}^3$, too.

(ii) Call a *line* $\mathcal{E} \subset \mathbb{R}^3$ *large* if it contains at least 3 elements of \mathcal{H} .

A *plane* $\mathcal{S} \subset \mathbb{R}^3$ is called **large** if $|\mathcal{S} \cap \mathcal{H}| \geq 4$. A large plane is **huge** if it contains at least 4 elements of \mathcal{H} , none 3 of them is colinear (i.e. an \mathcal{S} contains a 4 -element simplex of \mathcal{H}). \square

Of course, when we talk about *lines* and *planes* we consider the points of \mathcal{H} only.

Now we start the proof of our main result (which is an equivalent reformulation of Conjecture 4):

Theorem 16 *Assuming $\mathcal{H} \subset \mathbb{R}^3$, $|\mathcal{H}| = m \geq 24$ and \mathcal{H} spans \mathbb{R}^3 (i.e. \mathcal{H} is not coplanar) we have*

$$\text{simp}(\mathcal{H}) \geq \binom{\lfloor m/2 \rfloor}{3} + \binom{\lceil m/2 \rceil}{3} \quad (7)$$

and equality holds just in the case if \mathcal{H} is contained in two skew (detour) lines and these lines contain $\lfloor m/2 \rfloor$ and $\lceil m/2 \rceil$ many points of \mathcal{H} .

The smaller cases $4 \leq |\mathcal{H}| \leq 23$ will be described in <http://math.uni-pannon.hu/~szalkai/simplexes.html>.

Proof The case when \mathcal{H} is contained in two skew lines is trivial:

$$\text{simp}(\mathcal{H}) \geq \binom{\lfloor m/2 \rfloor}{3} + \binom{\lceil m/2 \rceil}{3} \text{ in this case (see Lemma 20).}$$

The only thing left is to prove that *all other cases* contain more simplexes.

The rest of the paper is devoted to the proof. We investigate two cases in the forthcoming two subsections: does \mathcal{H} contain a 5 -element simplex or not. \square

3.1 The CASE “ \mathcal{H} contains a 5-element simplex”

Let

$$\mathcal{R}_* := \{R_1, R_2, R_3, R_4, R_5\} \subseteq \mathcal{H}$$

be a fixed 5 -element simplex in \mathcal{H} .

Definition 17 Call any triplet $\{Q_1, Q_2, Q_3\} \subseteq \mathcal{H}$ **extendable** iff

- either $\{Q_1, Q_2, Q_3\}$ is a (3 -element) simplex
- or $\{Q_1, Q_2, Q_3, R_i\}$ is a (4 -element) simplex for some $R_i \in \mathcal{R}_*$
- or $\{Q_1, Q_2, Q_3, R_i, R_j\}$ is a (5 -element) simplex for some $R_i, R_j \in \mathcal{R}_*$. \square

We will explore that the number of extendable triplets—and so $\text{simp}(\mathcal{H})$ - is greater than $\binom{\lfloor m/2 \rfloor}{3} + \binom{\lceil m/2 \rceil}{3}$ in (7).

Lemma 18 *The incidencymatrices of non-extendable triplets (up to permutations) are only:*

Q1	Q2	Q3	R1	R2	R3	R4	R5	
(a)								
0	1	1	0	0	0	1	1	Line 1
1	0	1	1	1	1	0	0	Plane 1
1	1	1	0	0	0	1	1	Plane 2

Q1	Q2	Q3	R1	R2	R3	R4	R5	
(b)								
1	0	1	0	0	0	0	1	Line 1
0	1	1	0	0	0	1	0	Line 2
1	1	0	1	1	1	0	0	Plane 1
1	1	1	0	0	0	1	1	Plane 2

and

Q1	Q2	Q3	R1	R2	R3	R4	R5	
(c)								
0	1	1	0	0	0	0	1	Line 1
1	0	1	0	0	1	1	0	Line 2
1	1	0	1	1	0	0	0	Plane 1
1	1	1	0	0	1	1	1	Plane 2

Proof So we are given the *non*-extendable triplet $\{Q_1, Q_2, Q_3\} \subseteq \mathcal{H}$. This means that $\{Q_1, Q_2, Q_3\}$ is *not* colinear and for each $R_j \in \mathcal{R}_*$ we have either $R_j \notin [Q_1, Q_2, Q_3]$ or $R_j \in [Q_u, Q_v]$ for some $u, v \in \{1, 2, 3\}$. Moreover, none of the quintlets $\{Q_1, Q_2, Q_3, R_{j_1}, R_{j_2}\}$ is a 5 -element simplex for all $1 \leq j_1, j_2 \leq 5$. So **either** $\{R_{j_1}, R_{j_2}, Q_u\}$ is colinear, **or**

$$\{R_{j_1}, R_{j_2}, Q_u, Q_v\} \text{ is coplanar for some } u \text{ and } v. \tag{8}$$

(Since the previous case implies the last one, we have to deal only with the last one.)

Sublemma For some permutation (a, b, c, d, e) of the set $(1, 2, 3, 4, 5)$ and for some $1 \leq u < v \leq 3$ the points $[R_a, R_b, R_c, Q_u, Q_v]$ and $[R_d, R_e, Q_u, Q_v]$ are coplanar, further the line $[Q_u, Q_v]$ may contain at most one R_i point.

Proof All the $\binom{5}{2} = 10$ pairs $\{R_{j_1}, R_{j_2}\}$ are coplanar with some pair $\{Q_u, Q_v\}$ but the total number of pairs $\{Q_u, Q_v\}$ is only $\binom{3}{2} = 3$. By the pigeonhole principle at least 4 pairs of kind $\{R_{j_1}, R_{j_2}\}$ are coplanar with the *same* $\{Q_u, Q_v\}$. The set of these 4 pairs of $\{R_{j_1}, R_{j_2}\}$ can not be disjoint. If e.g. $\{R_1, R_2\}$ and $\{R_2, R_3\}$ are both coplanar with $\{Q_u, Q_v\}$ then clearly all the five points $\{R_1, R_2, R_3, Q_u, Q_v\}$ must be coplanar. Since all the points $\{R_1, R_2, R_3, R_4, R_5\}$ may not lie on a plane, the only possibility is that we have the two planes $[R_a, R_b, R_c, Q_u, Q_v]$ and $[R_d, R_e, Q_u, Q_v]$ (up to isomorphism).

If (at least) two points of $\{R_1, R_2, R_3, R_4, R_5\}$ are incident on the line $[Q_u, Q_v]$ then one of the planes would contain at least *four* points of them, contradicting to the assumption that $\{R_1, R_2, R_3, R_4, R_5\}$ is a 5 -element simplex. \square

Denote the number of the $R_j \in \mathcal{R}_*$ points incident on the plane $[Q_1, Q_2, Q_3]$ by db , clearly $0 \leq db \leq 3$.² In each of the cases below we cite Sublemma without mentioning it. (The indices of the points R_i follow the numbering of the Tables above—the result of our computer programs.)

Case $db = 0$: We have two planes $[R_1, R_2, R_3, Q_1, Q_2]$ and $[R_4, R_5, Q_1, Q_2]$. We will show that one of the six quintlets $\{Q_1, Q_2, Q_3, R_i, R_j\}$ ($j = 4, 5, i = 1, 2, 3$) is a 5 -element simplex. If one of them, “coded” by (i, j) is *not* a 5 -element simplex then (iff) *either* $[Q_3, R_j, R_i]$ forms a line *or* $[Q_t, Q_3, R_j, R_i]$ forms a plane either for $t = 1$ or for $t = 2$ (using $db = 0$ again), the latter case extends the first one, too. Can Q_3 be included in all of these *twelve* planes? In fact, we have at most four planes $[Q_t, Q_3, R_j]$ only. Consider first R_1 . If both $\{Q_1, Q_2, Q_3, R_4, R_1\}$ and $\{Q_1, Q_2, Q_3, R_5, R_1\}$ are *not* simplexes, then

$$\begin{aligned} & \text{(either } R_1 \in [Q_1, Q_3, R_4] \text{ or } R_1 \in [Q_2, Q_3, R_4]) \\ & \text{and (either } R_1 \in [Q_1, Q_3, R_5] \text{ or } R_1 \in [Q_2, Q_3, R_5]). \end{aligned}$$

Using $db < 2$ we must have

$$R_1 \in [Q_1, Q_3, R_4] \cap [Q_1, Q_3, R_5] \text{ OR } R_1 \in [Q_2, Q_3, R_4] \cap [Q_2, Q_3, R_5],$$

that is $R_1 \in [Q_1, Q_3, R_4, R_5]$ OR $R_1 \in [Q_2, Q_3, R_4, R_5]$.

These cases are equivalent to “ $[Q_1, R_4, R_5]$ is *colinear*” and “ $[Q_2, R_4, R_5]$ is *colinear*”, respectively. These two cases exclude each other, let us suppose that $[Q_1, R_4, R_5]$ is *colinear*. The same holds for R_2 and R_3 , so $R_1, R_2, R_3 \in [Q_1, Q_3, R_4, R_5]$. Together with $R_1, R_2, R_3 \in [R_1, R_2, R_3, Q_1, Q_2]$ but $Q_3, R_4, R_5 \notin [R_1, R_2, R_3, Q_1, Q_2]$ we must conclude that R_1, R_2, R_3 are *colinear*—a contradiction. This argument justifies that one of the quintlets $\{Q_1, Q_2, Q_3, R_j, R_i\}$ is a 5 -element simplex.

Case $db = 1$: Let $R_1 \in [Q_1, Q_2, Q_3]$ which also implies e.g. that $R_1 \in [Q_1, Q_2]$. By Sublemma we must have two planes $[R_a, R_b, R_c, Q_u, Q_v]$ and $[R_d, R_e, Q_u, Q_v]$, but what can (u, v) be?

If $(u, v) = (1, 2)$ then we have the planes $[R_1, R_2, R_3, Q_1, Q_2]$ and $[R_1, R_4, R_5, Q_1, Q_2]$ using the Sublemma and $R_1 \in [Q_1, Q_2]$. Now each of the *four* pairs $\{R_2, R_4\}, \{R_2, R_5\}, \{R_3, R_4\}, \{R_3, R_5\}$ must be coplanar either with $\{Q_3, Q_1\}$ or with $\{Q_3, Q_2\}$. In any case the pigeonhole principle implies, e.g. that $[R_2, R_4, R_5, Q_3, Q_1]$ and $[R_3, R_4, R_5, Q_3, Q_2]$ are new planes which have 3 common points—a contradiction.

If $(u, v) \neq (1, 2)$, say $(u, v) = (1, 3)$ then suppose that the points $[R_2, R_3, R_4, Q_1, Q_3]$ and $[Q_1, Q_2, R_1, R_5]$ and $[Q_1, Q_2, Q_3, R_1]$ are coplanar while $R_5 \notin [R_2, R_3, R_4, Q_1, Q_3] \cup [Q_1, Q_2, Q_3, R_1]$.

² The cases (a) and (b) in the statement belong to $db = 2$ while case (c) belongs to $db = 3$.

We now show that at least one of the *three* quintlets $\{Q_1, Q_2, Q_3, R_i, R_5\}$ ($i = 2, 3, 4$) is a 5 -element simplex. If not, then either $[Q_2, R_5, R_i]$ forms a line or $[Q_t, Q_2, R_5, R_i]$ forms a plane for some $i \in \{2, 3, 4\}$ and $t \in \{1, 2\}$ (using $db < 2$), the latter case extends the first one, too. By the pigeonhole principle there is a t for which there are (at least) *two* i such that $[Q_t, Q_2, R_5, R_i]$ forms a plane, say $[Q_3, Q_2, R_5, R_2], [Q_3, Q_2, R_5, R_4]$ and $[Q_1, Q_2, R_5, R_3]$. But in this case $\{R_5, R_2\}$ can *not* be coplanar with any $\{Q_r, Q_s\}$ which implies that the quintlet $\{Q_1, Q_2, Q_3, R_5, R_2\}$ is an 5 -element simplex!

Case $db = 2$: Using $R_4, R_5 \in [Q_1, Q_2, Q_2]$ we have two Subcases:

Subcase (i): $R_4, R_5 \in [Q_2, Q_3]$ which implies that the two planes mentioned in Sublemma must e.g. be $[R_1, R_2, R_3, Q_1, Q_3]$ and $[R_4, R_5, Q_1, Q_3] = [R_4, R_5, Q_1, Q_2, Q_3]$.

This is Table (a) in the present Lemma.

Subcase (ii): $R_4 \in [Q_2, Q_3]$ and $R_5 \in [Q_1, Q_3]$ implies that this two planes must be $[R_1, R_2, R_3, Q_1, Q_2]$ and $[R_4, R_5, Q_1, Q_2] = [R_4, R_5, Q_1, Q_2, Q_3]$ since $db = 2$, which is Table (b).

Case $db = 3$: Assuming *three* $R_i \in [Q_1, Q_2, Q_3]$ we again have two Subcases:

Subcase (i): $R_3 \in [Q_1, Q_2], R_1 \in [Q_1, Q_3]$ and $R_2 \in [Q_2, Q_3]$. Sublemma gives us the planes (by symmetry) $[R_3, R_4, R_5, Q_1, Q_2]$ and $[R_1, R_2, R_3, Q_1, Q_2, Q_3]$.

Consider now the pairs $\{R_1, R_4\}, \{R_1, R_5\}, \{R_2, R_4\}$ and $\{R_2, R_5\}$ —using (8) we get pairs $\{Q_u, Q_v\}$ such that the points $[R_i, R_j, Q_u, Q_v]$ are coplanar. We must have $\{u, v\} \neq \{1, 2\}$ since there are no three colinear points among $\{R_1, \dots, R_5\}$. Only *two* possibilities for $\{u, v\}$ remained for the *four* pairs $\{R_i, R_j\}$ we are considering, so at least two of them shares the same $\{u, v\}$.

Suppose e.g. that $[R_1, R_4, Q_1, Q_3]$ and $[R_1, R_5, Q_1, Q_3]$ are coplanar which means

$$R_4, R_5 \in [R_1, R_4, R_5, Q_1, Q_3] \cap [R_3, R_4, R_5, Q_1, Q_2].$$

Since *no* four point among $\{R_1, \dots, R_5\}$ are coplanar, the points $[R_2, R_4, Q_2, Q_3]$ and $[R_2, R_5, Q_2, Q_3]$ are coplanar, too, which implies that

$$R_4, R_5 \in [R_2, R_4, R_5, Q_2, Q_3].$$

The last two relation implies

$$R_4, R_5 \in [R_4, R_5, Q_1, Q_3] \cap [R_4, R_5, Q_1, Q_2] \cap [R_4, R_5, Q_2, Q_3]$$

which is a *contradiction* since $[Q_1, Q_2, Q_3]$ are coplanar but not colinear and $R_4, R_5 \notin [Q_1, Q_2, Q_3]$ and $R_4 \neq R_5$.

Subcase (ii): $R_3, R_4 \in [Q_1, Q_3]$ and $R_5 \in [Q_2, Q_3]$. By Sublemma we must have one more large plane *either* $[R_1, R_2, Q_2, Q_3]$ *or* $[R_1, R_2, Q_1, Q_2]$ since the plane $[R_1, R_2, Q_1, Q_3, R_3, R_4]$ is impossible.

If $[R_1, R_2, Q_2, Q_3]$ is a plane, consider the pairs $\{R_1, R_4\}, \{R_1, R_3\}, \{R_2, R_4\}$ and $\{R_2, R_3\}$. Using (8) we get pairs $\{Q_u, Q_v\}$ such that the points $[R_i, R_j, Q_u, Q_v]$ are

coplanar. Since $R_3, R_4 \in [Q_1, Q_3]$ and $\{Q_1, Q_2, Q_3\}$ are not colinear we must have $\{Q_u, Q_v\} = \{Q_1, Q_3\}$ which implies that $[R_1, R_2, R_3, R_4]$ would be coplanar—a contradiction.

The plane $[R_1, R_2, Q_1, Q_2]$ implies that we are talking about Table (c).

End of the proof of Lemma 18. □

Lemma 19 Any non-extendable triplet $\{Q_1, Q_2, Q_3\}$ is determined by (at most) two of its elements.

Proof We have to discuss the three cases from the previous lemma.

Case (a) $Q_1 \notin [R_4, R_5]$ since $\{Q_1, Q_2, Q_3\}$ is not colinear (see Table a), line 1). So we may let

$$line2 := [Q_1, Q_3] = Plane1 \cap Plane2 = [R_1, R_2, R_3] \cap [R_4, R_5, Q_1],$$

so $Q_3 = line1 \cap line2 = [R_4, R_5] \cap line2$ which means that Q_3 is determined by Q_1 . Since $Q_2 \in [R_4, R_5]$ is arbitrary, the triplet $\{Q_1, Q_2, Q_3\}$ is determined by Q_1 and Q_2 in this case. (Moreover $Q_1 \in [R_1, R_2, R_3]$ must be assumed.)

Case (b) $Q_3 \notin [R_4, R_5]$ since $\{Q_1, Q_2, Q_3\}$ is not colinear (see Table (b), line 1). So we may let

$$line3 := [Q_1, Q_2] = Plane1 \cap Plane2 = [R_1, R_2, R_3] \cap [R_4, R_5, Q_3],$$

so $Q_1 = line1 \cap line3 = [Q_3, R_5] \cap line3$

and $Q_2 = line2 \cap line3 = [Q_3, R_4] \cap line3$ which mean that Q_3 determines Q_1 and Q_2 in this case.

Case (c) If e.g. $Q_2 \notin [R_1, R_2]$ (or $Q_1 \notin [R_1, R_2]$) then we let

$$line3 := [Q_1, Q_2] = Plane1 \cap Plane2 = [R_1, R_2, Q_2] \cap [R_3, R_4, R_5],$$

so $Q_1 = line2 \cap line3 = [R_3, R_4] \cap line3$

and $Q_3 = line2 \cap line1 = [R_3, R_4] \cap [Q_2, R_5]$

which mean that Q_2 determines Q_1 and Q_3 .

Clearly either $Q_1 \notin [R_1, R_2]$ or $Q_2 \notin [R_1, R_2]$ must hold since otherwise $[R_1, R_2] = [Q_1, Q_2] \subset [R_3, R_4, R_5]$ would hold but $[R_1, R_2, R_3, R_4, R_5]$ is a contradiction.

The subcase $[R_3, R_4] \cap [Q_2, R_5]$ means that the points $\{R_1, R_2, R_3, R_4, R_5, Q_1, Q_2, Q_3\}$ do not satisfy any case of Lemma 18 which means that $\{Q_1, Q_2, Q_3\}$ is an extendable triplet. □

By the above result we have that the number of the non-extendable triplets is at most $3 \cdot \binom{m-5}{2}$, so we have at least $\binom{m-5}{3} - 3 \cdot \binom{m-5}{2}$ many extendable triplets. This implies

$$simp(\mathcal{H}) \geq 1 + \binom{m-5}{3} - 3 \cdot \binom{m-5}{2}$$

which is larger than $\binom{\lfloor m/2 \rfloor}{3} + \binom{\lceil m/2 \rceil}{3}$ in (7) since the corresponding cubic function

$$1 + \binom{m-5}{3} - 3 \cdot \binom{m-5}{2} - 2 \binom{m/2}{3} > 0$$

has the roots $m_1 \approx 4.747$, $m_2 \approx 5.637$ and $m_3 \approx 23.615$.

End of the proof of CASE “ \mathcal{H} contains a 5-element simplex”. □

3.2 The CASE “ \mathcal{H} does not contain a 5-element simplex”

In Lemmae 20 through 22 we do not use the assumption that \mathcal{H} does not contain a 5-element simplex. The proof itself of the Case “ \mathcal{H} does not contain a 5-element simplex” starts after Lemma 22

Lemma 20 *If two skew lines contain all points of \mathcal{H} then $\text{simp}(\mathcal{H})$ is minimal if and only if these lines contain $\lfloor \frac{m}{2} \rfloor$ and $\lceil \frac{m}{2} \rceil$ points, respectively.*

Proof Let these lines contain x and $m - x$ points ($m \geq 4$), then we have only 3 -element simplexes:

$$\begin{aligned} \text{simp}(\mathcal{H}) &= \binom{x}{3} + \binom{m-x}{3} \\ &= \frac{m-2}{2} \cdot x^2 - \frac{m(m-2)}{2} \cdot x + \binom{m}{3} \\ &\geq 2 \cdot \binom{m/2}{3} \end{aligned} \tag{9}$$

□

Lemma 21 *If $\text{simp}(\mathcal{H})$ is minimal then there is a large plane in \mathcal{H} .*

Proof If not, then there are no large lines as well. But then each 5 points of \mathcal{H} forms a simplex, which means that $\text{simp}(\mathcal{H})$ is maximal by [1]. □

Lemma 22 *If $|\mathcal{H}| \geq 8$ and \mathcal{H} but a single point P_0 is contained in a single plane \mathcal{S} then $\text{simp}(\mathcal{H})$ is not minimal.*

Proof In this case

$$\begin{aligned} \text{simp}(\mathcal{H}) &= \text{simp}_3(\mathcal{H}) + \text{simp}_4(\mathcal{H}) + \text{simp}_5(\mathcal{H}) \\ &= \text{simp}(\mathcal{S}) + 0 \\ &\geq \binom{m-3}{3} + 1 + \binom{m-4}{2} \end{aligned} \tag{10}$$

using $m \geq 8$ and the results in [16].

It is easy to see that $\binom{m-3}{3} + 1 + \binom{m-4}{2} > 2 \binom{m/2}{3}$ for each $m \in \mathbb{N}$ since the corresponding cubic equality has roots $m_1 \approx -0.296$, $m_{2,3} \approx 5.148 \pm -0.738i$. □

Proof of Theorem 16 CASE “there is no 5-element simplex in H ” assuming $|H| \geq 8$.

Since \mathcal{H} spans \mathbb{R}^3 there is a tetrahedron $\{A_1, A_2, A_3, A_4\} \subset \mathcal{H}$.
Supposing that no 5 -element simplex is contained in \mathcal{H} we know that all other points

of \mathcal{H} must lie on the infinite surface of this tetrahedron (on the four planes determined by these four points).

Case I. There are 4 points in general position (i.e. no 3 is lying in a line) on (at least) one of the planes—i.e. this is a huge plane. Let us suppose this plane $[A_2, A_3, A_4]$ with $A_5 \in [A_2, A_3, A_4]$ where A_2, A_3, A_4, A_5 are in general positions.

Now consider all the four tetrahedrons with common vertex A_1 , that is $\{A_1, A_2, A_3, A_4\}$, $\{A_1, A_2, A_3, A_5\}$, $\{A_1, A_2, A_4, A_5\}$ and $\{A_1, A_3, A_4, A_5\}$. \mathcal{H} must be contained in the intersection of their surfaces. This intersection is the plane $[A_2, A_3, A_4, A_5]$ plus the (possible) lines $[A_1, C_1]$, $[A_1, C_2]$ and $[A_1, C_3]$ where C_1, C_2, C_3 are the following points of the plane $[A_2, A_3, A_4, A_5]$:

$C_1 := [A_2, A_5] \cap [A_3, A_4]$, $C_2 := [A_4, A_5] \cap [A_2, A_3]$ and $C_3 := [A_3, A_5] \cap [A_2, A_4]$.³

Subcase 1 $\mathcal{H} \setminus \{A_1\}$ is contained in the plane $[A_2, A_3, A_4]$. This case was discussed in Lemma 22.

Subcase 2 There is at least two points of \mathcal{H} outside of the plane $[A_2, A_3, A_4]$ (one of them is A_1). By symmetry we may suppose that the new point $D \in \mathcal{H}$ is incident on $[A_1, C_1]$, say $D \in [A_1, C_1]$, $D \neq C_1$. Applying the argument before Subcase 1 to D instead of A_1 we may conclude that *no* elements of \mathcal{H} may lie on the lines $[A_1, C_2]$ and $[A_1, C_3]$.

In other words: \mathcal{H} is contained in the plane $[A_2, A_3, A_4]$ plus the line $[A_1, C_1]$.

Subsubcase (2.a) There is a further point $E \in \mathcal{H}$ on the plane $[A_2, A_3, A_4]$, $E \neq C_1$. We may choose two (fixed) indices $i, j \in \{2, 3, 4, 5\}$ so that $C_1, E \notin [A_i, A_j]$ which implies that $\{A_1, A_i, A_j, D, E\}$ is a 5 - *element* simplex - a contradiction.

Subsubcase (2.b) No further points of \mathcal{H} are on the plane $[A_2, A_3, A_4]$ just $\{A_2, A_3, A_4, A_5\}$, i.e. $\mathcal{H} \setminus \{A_2, A_3, A_4, A_5\} \subset [A_1, C_1]$.

In the subsubsubcase $C_1 \in \mathcal{H}$ we can find 3 -element simplexes are on the line $[A_1, C_1]$ while $\{A_2, A_3, A_4, A_5\}$ and $\{A_1, A_i, A_j, X\}$ are 4 -element simplexes for all $X \in [A_1, C_1]$, so

$$\text{simp}(\mathcal{H}) \geq \binom{m-4}{3} + 1 + \binom{m-5}{2} \cdot \binom{4}{2} > 2 \cdot \binom{m/2}{3}$$

while in the subsubsubcase $C_1 \notin \mathcal{H}$ we have

$$\text{simp}(\mathcal{H}) \geq \binom{m-4}{3} + 1 + \binom{m-4}{2} \cdot \binom{4}{2} > 2 \cdot \binom{m/2}{3}$$

for $m \geq 7$.

Case II.: there are no 4 points in general position on any of the planes of the tetrahedron $\{A_1, A_2, A_3, A_4\}$. Now let us say the plane $[A_1, A_2, A_3]$ contains at least 4 points with the line $[A_1, A_2]$ containing all (at least 3 but A_3) points of this plane. Then, the plane $[A_1, A_2, A_4]$ also must contain no points outside the line $[A_1, A_2]$

³ Some of the points C_1, C_2, C_3 may not exist when $[A_1, C_i]$ is parallel to $[A_2, A_3, A_4, A_5]$. However, at least one of the points C_1, C_2, C_3 must exist, say C_1 does. Further, C_1, C_2, C_3 need not belong to \mathcal{H} , this question is irrelevant until Subsubcase (2.b).

except A_3 . This means that no edges (as infinite lines) of the tetrahedron may contain more than 2 points except $[A_1, A_2]$ and $[A_3, A_4]$. This case of two skew lines was calculated in Lemma 20.

End of the Proof of Theorem 16.

□

References

1. I. Szalkai, C. Laflamme, Counting Simplexes in \mathbb{R}^n . Hung. J. Ind. Chem. **23**, 237–240 (1995). HU ISSN 0133-0276, ref: Chemical Abstract **123**, 290982 (1995)
2. M. Hujter, *On the Definition of the n-Dimensional Regular Simplexes Using Inequalities* (in Hungarian), Manuscript 6 Mar 2009, see. <http://math.bme.hu/~hujter/dekatetr.pdf>
3. I. Szalkai, Handling multicomponent systems in \mathbb{R}^n , I Theoretical results. J. Math. Chem. **25**, 31–46 (1999)
4. S. Kumar, Á. Pethő, Note on a combinatorial problem for the stoichiometry of chemical reactions. Intern. Chem. Eng. **25**, 767–769 (1985)
5. Á. Pethő, The linear relationship between stoichiometry and dimensional analysis. Chem. Eng. Technol. **13**, 328–332 (1990)
6. L.T. Fan, B. Bertók, F. Friedler, S. Shafie, Mechanisms of ammonia-synthesis reaction revisited with the aid of a novel graph-theoretic method. Hung. J. Ind. Chem. **29**, 71–80 (2001)
7. L.T. Fan, B. Bertók, F. Friedler, A graph-theoretic method to identify candidate mechanisms. Comput. Chem. **26**, 265–292 (2002)
8. P. Sellers, *Postscript: On the Algebra, Combinatorics & Algorithms for Chemical Reaction Networks*. manuscript
9. P. Sellers, *Torsion in Biochemical Reaction Networks*. Bulletin of Mathematical Biology (Accepted)
10. I. Szalkai, Generating minimal reactions in stoichiometry using linear algebra. Hung. J. Ind. Chem. **19**, 289–292 (1991)
11. Á. Pethő, *Lectures on Linear Algebraic Methods in Chemical Engineering Mathematics (stoichiometry)*. Preprint 1990 (Institut für Technische Chemie, Univ. Hannover, Germany)
12. Á. Pethő, On a class of solutions of algebraic homogeneous linear equations. Acta Math. Hung. **18**, 19–23 (1967)
13. J. Appel, M. Otarod, P.H. Sellers, Mechanistic study of chemical reaction systems. Ind. Eng. Chem. Res. **29**, 1057–1067 (1990)
14. I. Szalkai, C. Laflamme, Gy. Dósa, On the maximal and minimal number of bases and simple circuits in matroids and the extremal constructions. *Pure Math. Appl. (PUMA)* **15**, 383–392, HU ISSN 1218–4586 (2006)
15. I. Szalkai, *Private Communication to Cl.Laflamme*, 24 Feb 2000
16. I. Szalkai, C. Laflamme, Counting Simplexes in \mathbb{R}^3 . *Electr. J. Comb.* **5**(1), Res. Paper 40, 11 pp (1998). <http://www.combinatorics.org>. Printed version: *J. Combin.* **5**, 597–607, ref: Current Math. Publ. 1998, No. 16., ZBL 902.05076 (1998)
17. P. Sellers, *Private Communication* 7 May 2007