

Earth's biodiversity as interpreted by a chemist

Jan Turulski*

16-035 Czarna Wies Koscielna, Poland, and Chemistry Institute, University at Bialystok, 15-443 Bialystok, Poland

Jan Niedzielski

Chemistry Department, Warsaw University, 02-089 Warsaw, Poland
E-mail: niedziel@chem.uw.edu.pl

Received 10 November 2003; revised 8 March 2004

Exotic states such as quasicrystals are postulated to exist apart from the normal states such as crystals or liquids in complex systems containing at least three independent components. The number of exotic states increases asymptotically as the number of components increases. Assuming that the biosphere incorporates about 30 elements and assuming the occurrence of all of the thermodynamically allowed processes the number of one-phase exotic states can, intuitively, be used as a measure of Earth's biodiversity. A structure of links between the individual exotic systems is similar to that between species in the genealogical chart of biosphere.

KEY WORDS: biodiversity, graph theory, thermodynamic equilibrium

AMS subject classification: 80A50, 92E10, 92B10

1. Introduction

Species diversity on Earth is a fundamental problem in biology. For many different reasons a number of species that lived on Earth during one or another geological period cannot be estimated readily. A current estimate is also very approximate; it is argued that there are now from 10 up to 80 [1,2] or even 100 millions of different species of living organisms [3]. Two opposing processes determine this number: evolutionary productivity of nature and natural or catastrophic extinctions. In some geological periods the latter process resulted in disappearance of as many as 95% of the then existing species [4].

Some authors estimate that during the whole history of Earth the rate of extinction was about one species a year. Using these numbers, a rough estimation of the number of taxonomic groups, such as for instance genera or spe-

*Corresponding author.

E-mail: turul@uwb.edu.pl

cies, that ever lived on Earth can be attempted. This number seems to be on the order of billions. It can be argued that the total biodiversity should be one way or another connected with the diversity of some chemical systems. Obviously, every living organism can be treated as, in fact, an open system, remote from equilibrium, and spacially and temporally ordered [5]. However, thus far neither biologists nor chemists took interest in the diversity of such chemical systems that could underlie living organisms. In physics and chemistry the situation is diametrically different. Owing to the common effort a link has been established between the physical diversity of elementary particles (protons, neutrons and electrons) and the chemical diversity represented by chemical elements.

The aim of this work is to indicate a correspondence between diversity of some chemical systems and species diversity on Earth.

2. Theory

In thermodynamics, a state equilibrium of the system is described by a certain function G , known as thermodynamic potential. In the systems with C independent components

$$G = G(p, T, c_1, c_2, \dots, c_C), \quad (1)$$

where p , T , c_1 , c_2, \dots, c_C denote respectively the pressure, the temperature and the concentration of individual independent components.

Individual types of states of the complex systems are known as phases, and are described by the individual types of the state functions. A condition for thermodynamic equilibrium in the systems with C independent components and P phases is equality of chemical potentials, $\mu_i^{(j)}$, for the i th component ($i = 1, 2, \dots, C$) in the individual phases ($j = 1, 2, \dots, P$) [6, 7]:

$$\mu_i^{(1)} = \mu_i^{(2)} = \dots = \mu_i^{(P)}, \quad (2)$$

where $i = 1, 2, \dots, C$.

This set of equations leads to the well-known Gibbs phase rule that describes the connection between C , P and the number of degrees of freedom, f .

$$f = C - P + 2. \quad (3)$$

Many authors [8–13] have noticed and offered different interpretations for a striking similarity between equation (2) and a famous Euler's formula that gives a fundamental relationship between vertices, V , edges, E , and faces, F , of a planar graph on the surface of two-dimensional sphere, henceforth called 2-sphere.

$$F = E - V + 2. \quad (4)$$

We have recently shown [14] that equations (3) and (4) are so similar, because there is such a vectorial function $R = (X, Y)$, dependent on thermodynamic parameters, whose picture can be visualized as a certain surface isomorphic with 2-sphere. The thermodynamic potential of this surface can be written in the form:

$$G = G(A(R_i, R_k), B(R_j, R_l), \dots, C(R_{f-1}, R_f)), \quad (5)$$

where R_x is the point of the 2-sphere surface. A number of internal functions A, B, \dots in the G potential is equal to the number of independent components, C .

It is seen that G defined in equation (5) can be represented by a planar graph on the R surface, called by us a graph of state [14]. The edges in the graph of state are the geodesic lines that link the individual vertices, R_i and R_k . The edges represent the individual independent components of the system. The sequence of edges closing certain area on the 2-sphere surface creates a face. This face represents a phase formed by those components that correspond to the edges enclosed in this sequence.

An example for the equilibrium graph of state G for the five-component (independent components A, B, C, D, E) and three-phase system (phases α, β, γ) is shown in figure 1(a). Equilibrium phase transition without formation of new phases is described by the change in the lengths of the geodesic lines that link the individual vertices. Phase transition transforms G graph into G' graph and is described by gluing of two graph G vertices. The phase composition is determined by the edges that form a given face. Equilibrium phase transition without formation or disappearance of phases is described by the change in the lengths of the geodesic lines that link the individual vertices.

The change in thermodynamic state is accompanied by G being transformed into G' which is brought about by gluing or disconnecting of some graph G vertices.

All possible graphs of state and their number for the one-phase systems with C from 1 up to 5 are shown in figure 1(b). Note that for the systems composed from more than two constituents there exist more than one graph of state. For example, at $C = 2$ $\eta = 2$ while at $C = 3$ $\eta = 3$. Whilst these graphs are topologically different, always one linear graph is available that is a usual generalization of the corresponding graph at $C = 1$ or $C = 2$. These are normal graphs. Only at $C \geq 3$ do the graphs described as a branched tree appear. We coined the name exotic graphs for such graphs. Consequently, the states represented by such graphs are also called exotic. In our recent paper [14] we argued that stable quasicrystals [15–18] can be represented by exotic graphs. Exotic state can be obtained from the normal one by phase transformation. Such a process is shown in figure 2 for a four-component system. Note that on the way from the normal to the exotic state the two-phase systems always occur (such

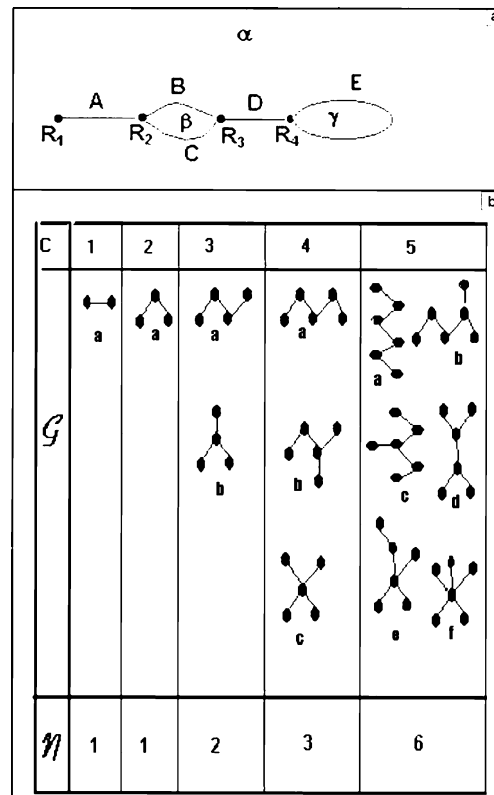


Figure 1. (a) Examples of the graphs of state for the three-phase (α , β , γ), five-component (A , B , C , D , E) system. Graph G is associated with the following distribution of components between the phases: $\alpha = (A, B, C, D, E)$; $\beta = (B, C)$; $\gamma = (E)$. (b) Graphs of state for the one-phase systems containing up to five independent components. η defines the number of graphs for the specific C . Graphs marked with "a" represent normal states; graphs marked with other letters represent exotic states.

two-phase systems are represented by the graphs in braces in figure 2). Note also that not every two-phase system undergoes transformation into the exotic state. It is seen in figure 2(a) that from among 11 available transformation channels, only five channels yield exotic state while six others return to the initial normal state. Thus examination of figure 2(a) reveals that under suitable conditions the normal G_n state transforms with a $4/11$ and $1/11$ probabilities into the exotic G_e and G'_e states, respectively. Similarly, under suitable conditions G_e undergoes transformation into another exotic state G'_e with a $2/13$ probability (see figure 2(b)).

Examination of the graphs of state for the systems $C > 4$ indicates that even for five independent components the G_e exotic states will appear that cannot be obtained directly from the normal G_n states as, for instance, the d and f phases

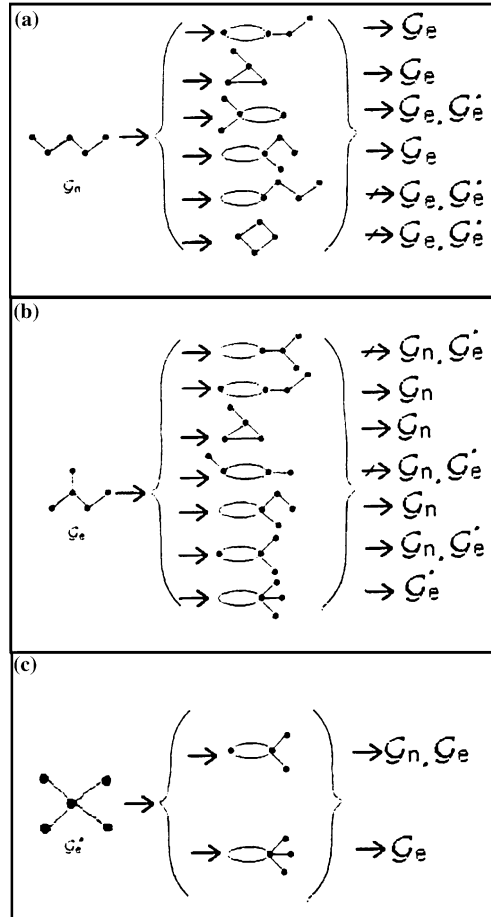
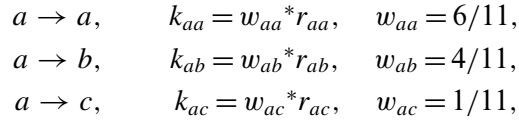


Figure 2. Phase transformations in the one-phase, four-component system.

in figure 1(b) at $C = 5$. Obviously, the exotic states of this type can descend from the less exotic states obtained earlier from the normal states. The f state in figure 1(b) can be formed from the phase represented by the d or e graph. It can be argued that the thermodynamical states in complex systems can be ordered in a diachronic series. A specific state will appear in a specific location in the series provided that it is preceded by the suitable ancestors.

Let us treat a phase transformation “one-phase system \rightarrow multiphase system” as a first-order kinetic process, characterized by the rate constant, k , that is a product of the probability, w , for the specific transformation channel, and the kinetic weight for this channel, r . We can suggest a mechanism for any possible phase transformation in a given thermodynamic system. For instance, a

mechanism of the transformations shown in figure 2(a) can be described by



where, according to figure 1b, a stands for the normal phase, G_n , while b and c stand for the exotic phases, G_e and G'_e . Solution of the differential kinetic equations for the rate of reaction described by this mechanism gives a genealogical tree linking individual phases in the complex system. The examples of the graphs representing genealogical tree for the system with four independent components are shown in figures 3(a) and (b) for two extreme situations. These graphs resulted from the solution of kinetic equations describing transformations shown in figure 2.

The example shown in figure 3(a) is based on the assumption that kinetic weights, r_{ij} , for all transformation channels are equal. In figure 3(b) such conditions are modified by assuming $r_{ac} = r_{ca} = 0$. Both types of graphs shown in Fig 3 are known to biological systematics. The process of formation of new species from two or greater number of different living organisms is called speciation as a result of allopolyploidy (for instance, the links between rape, *Brassica napus*, cauliflower, *Brassica oleracea*, and turnip, *Brassica campestris*). The second type, shown in figure 3b, is called cladogenesis.

The genealogical graph that represents the structure of links between all the graphs of states for one-phase four-components systems is shown in figure 3. Similar genealogical graphs can be conceived for more complex systems composed from larger numbers of independent constituents. If a plethora of exotic states is available, the genealogical graph will necessarily be more involved, retaining, however, a logical structure exhibited by the graph shown in figure 3. The numbers listed in Table 1 give the number of states for the systems with 10, 20 and 50 independent components [19]. Difficulties involved in the construction of genealogical graph are of course enormous. The number of graphs of state increases exponentially as C increases [21,22]. For every complicated system it is more convenient to replace C by the number of elements that form this system [22]. According to Zhao [23]

$$C = M + r' - Z, \quad (6)$$

where C is the number of constrained conditions other than Σ (mole fraction of the component) = 1, while r' is the number of independent reactions that do not occur due to kinetic restrictions. Z differs from zero if, for instance, the composition of one phase is the same as the composition of another one. If such special cases are ignored, $Z = 0$, as is assumed in our further considerations. Furthermore, we do not impose any kinetic restrictions on the reactions that occur in the system which is tantamount to the assumption that $r' = 0$. Then, irrespective

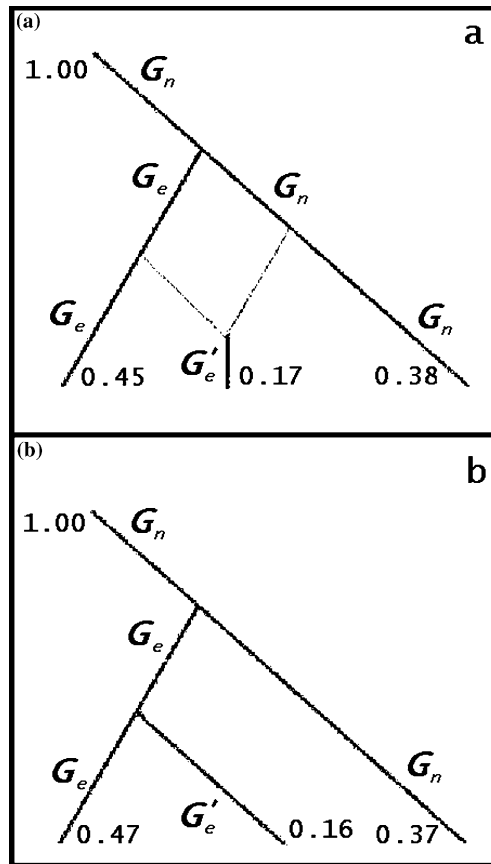


Figure 3. Genealogical graph for the states G_n , G_e and G'_e that describes the links between the states of the four-component system. figure 3(a) is obtained on the assumption that kinetic weights for all transformation channels are equal. figure 3(b) is obtained on the assumption that $r_{ac} = r_{ca} = 0$ while other r_{ij} weights are equal. The numbers at the bottom of graphs, besides the edges, give the fraction of the phase represented by the given edge that is formed after infinitely long time, from the system composed initially from the normal phase, G_n . (a) Transformation of the normal G_n graph into the other graphs. (b) Transformation of the exotic G_e graph into the other graphs. (c) Transformation of the exotic G'_e graph into the other graphs. Crossed arrows denote that the transformation does not occur.

of the number of components, the number of independent components is equal to M , the number of elements that form the individual components.

Living organisms are built from an enormous number of chemical compounds. Even rough estimation is of dubious value since in such a system as living organism a very notion of component is blurred. There are about 30 chemical elements that are involved in the tissue and functioning of living organisms [24]. A mean abundance, W , of chemical elements in a body of human weighing 70 kg is reported in Table 2. The presence of some elements, such as Al,

Table 1
 A number of different states, η , in equilibrium or close to equilibrium, in one-phase systems with a large number of constituents, C [19].

C	η
10	235
20	2144505
30	4.03×10^{10}
40	1.01×10^{15}
50	2.96×10^{19}

may be contested, since indispensability of aluminium for living organisms is by no means proven. Be as it may, the number of bioelements can be estimated as about 30 (electrons should also be taken into account since redox reactions certainly do occur in living organisms). If in a thermodynamic system bioelements are interlinked with all kinetically allowed reactions, then, according to Brinkley's equation (6), the number of independent components $C = 30$.

Inspection of Table 1 reveals that such equilibrium equivalents of living organisms can exist in 4.03×10^{10} types of state, as this number represents the chemical diversity of one-phase systems with 30 independent constituents. The logical structure of links between all of these states is analogous to that shown in figure 3. Even though the total diversity of biological species is very uncertain, it can reasonably be claimed that this number is close to the chemical diversity assessed above. It can also be claimed that a parent graph for individual states, that describes the links between the states in the chemical system, is qualitatively similar to the corresponding graphs known from taxonomy.

While the correspondence between non-equilibrium and equilibrium graphs for the system close to equilibrium seems to be well founded, it ceases to be so for the living organisms which are certainly very far from equilibrium. The application of the Gibbs' phase rule in the form of Eq. (2) may not be justified. For this reason, and it is not difficult to add others, our considerations on the links between biodiversity of life on Earth and the number of exotic equilibrium states are only speculative.

Some criticism can be rebutted readily. We do not think that exotic states of thermodynamic equilibria in some systems are tantamount to the states of biological systems that represent specific species of living organisms. States in equilibrium are different than the states far from equilibrium. All that we claim is the striking similarity between the number that represents the diversity of exotic equilibria states in some complex chemical systems and the number that represents the biodiversity of the Earth. We also draw attention to the similarity of relations between individual exotic states and between individual biological species. Thermodynamic parameters of living organisms and those of their

Table 2
Average abundance, W , of element, M , in a
human body weighing 70 kg [24].

M	W
As	18 mg
N	1.8 kg
B	?
Br	260 mg
Cl	95 g
Cr	?
Sn	14 mg
Zn	2.5 g
F	2.6 g
P	780 g
Al.	61 mg(?)
I	12-20 mg
Co	1.5 mg
Si	?
Mg	19 mg
Mn	12 mg
Cu	72 mg
Mo	?
Ni	?
K	140 g
Se	14 mg
S	140 g
Na	100 g
O	43 kg
V	0.11 mg
Ca	1 kg
C	16 kg
H	7 kg
Fe	4.2 g

surroundings are, of course, much different but we do not think that this makes our reasoning invalid. We do not try to define the state of the living organism. We are only interested in the number of feasible classes of living organisms. This number need not be determined by non-equilibrium thermodynamics but it can be determined using as a base equilibria in the systems constituted from about 30 bioelements assuming the occurrence of all kinetically allowed reactions. Of course, the same thermodynamic parameters that determine precisely the equilibrium state of the chemical system do not determine the state of the living organism. However, the states of living organisms do correlate with the states of exotic equilibria in some chemical systems. Analogous situation occurs in the world of atoms. The states of individual atoms certainly correlate with the states of corresponding sets of protons, neutrons and electrons, even though

the states of the set of elementary particles are determined by completely different parameters than is the state of the atom. The very notion of biological species has no clear-cut definition. Despite of this, we agree with Wilson [1] and with many others that this notion is useful. The same in chemistry—the chemical compound may not have a well defined composition but it does not preclude the use of this term.

Nevertheless, the search for the links between biology, as measured by the number of existing species, and chemistry, as measured by the number of thermodynamic states in complex systems, seems to be a worth while task. The striking similarity between the number of biological species and the number of equilibrium states for some chemical systems as well as the quantitative similarity between the genealogical graphs for both types of diversity may not be only accidental. Recall that graphs have been remarkably successful in enumerations in very disparate fields. Perhaps also the enumeration suggested in this work is more than playing with numbers. Perhaps the number of biological species, extinct and extant, is determined by the equilibrium thermodynamics of the system. Of course, properties of the individual biological species are determined by non-equilibrium thermodynamics. The analogy with the world of chemical atoms and elementary particles seems to be constructive. Of course, properties of the helium atom, ${}^4\text{He}$, are completely different from properties of the set of two protons, two neutrons and two electrons. Still, the helium atom, as well as any other atom, can be derived from the state of the suitable set of elementary particles. Thus, the diversity of chemical elements is measured by the diversity of the set of states for the proper number of protons, neutrons and electrons. In this work we transfer this reasoning on the correspondence between species diversity and diversity of equilibrium states for the system composed from 30 elements. Until hypotheses put on more firm ground are devised, speculation offered in this work is hoped to be of some interest.

3. Conclusions

Thermodynamic equilibrium states in complex systems have a similar structure of interlinks, a genealogical tree, as the structure of interlinks between living species in biology.

It is suggested that the diversity of thermodynamic equilibria states in the systems composed of 30 elements in which all kinetically allowed processes occur corresponds to the global biodiversity of Earth, estimated as the number of species, extinct and extant, that have ever lived on the planet.

The former diversity attains the number 4.03×10^{10} on the assumption that all kinetically allowed reactions occur. This number seems to be close to the number of biological species, extinct as well as extant, that can be created by the biosphere.

Acknowledgments

We are grateful to Professor J.R.E. Taylor, Biology Institute, University of Bialystok, for helpful comments.

Appendix

Equilibrium or non-equilibrium thermodynamic states of the complex system containing C independent constituents are described, in a given moment of time, by thermodynamic functions, G_i for thermodynamic parameters and their gradients, ∇ (in a simpler case by the differences in values for the pertinent parameters in the system and in its surroundings). If the subscript "i" enumerates all non-isomorphic states of the system, the system will be described by a set of non-isomorphic functions:

$$G_i = G_i(p, T, c_1, c_2, \dots, c_C; \nabla p, \nabla T, \nabla c_1, \nabla c_2, \dots, \nabla c_C) \quad i = 1, 2, \dots, \eta \quad (\text{A.1})$$

where η is the number of states.

For the systems composed from hundreds or thousands chemical components, N , we can expect G_i to be a function of monstrous number of arguments. It is not always the case. As shown in 1947 by Brinkley [23], if all kinetic reactions occur in the system then, irrespective of the complexity of the system, the number of independent components, C , can be replaced by the number of elements, M , from which the components of the system are formed.

$$C = M. \quad (\text{A.2})$$

However, even after this simplification η is not known generally. We have recently shown [15] that the problem simplifies considerably if only equilibrium thermodynamic states are taken into account. Under such conditions, the gradient arguments in thermodynamic functions become zero

$$G_i^e = G_i(p, T, c_1, c_2, \dots, c_C; 0, 0, \dots, 0) \quad i = 1, 2, \dots, \eta_e. \quad (\text{A.3})$$

Moreover, according to Gibbs phase rule, G_i^e can be given in the form:

$$G_i^e = G_i(R_1, R_2, R_3, \dots, R_f; 0, 0, \dots, 0) \quad i = 1, 2, \dots, \eta_e \quad (\text{A.4})$$

where f is the number of thermodynamic degrees of freedom for the system, and R_i is the coordinate of the point on a certain two-dimensional surface, Σ , isomorphic with the surface of sphere in the space of thermodynamic parameters.

$$\Sigma = [X(p, T, c_1, c_2, \dots, c_C), Y(p, T, c_1, c_2, \dots, c_C)]. \quad (\text{A.5})$$

If conditions resulting from the reference state for the system are imposed on the location for R_i [14], every R_i point will represent one rather than two degrees of freedom on the Σ surface. As shown in [14], the number of one-phase equilibrium states, η_e , is equal to the number of graphs that are the trees with C edges. For $C = 30$, according to [13], $\eta_e = 4.03 \times 10^{10}$.

$$\eta = \eta_e. \quad (\text{A.6})$$

Note that the equality is valid only if either some non-equilibrium states do not undergo degeneration into one equilibrium, state attaining equilibrium. or, even if they do, which is more likely, the contribution of the states undergoing degeneration can be ignored compared with the number of all states.

References

- [1] E.O. Wilson, *The Diversity of Life* (W.W.Norton, New York, 1993).
- [2] J. Weiner, *Life and Biosphere Evolution* (PWN, Warsaw, 1999), pp. 283-302, (in Polish)
- [3] P.J. Bryant, *Biodiversity and Conservation* (A Hypertext Book at URL address: <http://darwin.bio.uci.edu/~sustain/bio65/>, 2000), Chapter 6.
- [4] S.S. Stanley, *Earth System History* (W.H.Freeman, New York, 1999).
- [5] I. Prigogine and G. Nicolis, *Self-Organization in Non-equilibrium Systems: From Dissipative Structures to Order Through Fluctuations* (J. Wiley, New York, 1977).
- [6] M. Hillert, *Phase Equilibria, Phase Diagrams and Phase Transformations* (Cambridge University Press, Cambridge, 1998).
- [7] I.M. Klotz and R.M. Rosenberg, *Chemical Thermodynamics* (J.Wiley, New York, 1994).
- [8] I. Levin, *J. Chem. Edu.* 23 (1946), 183.
- [9] A.L. Seifer and V.S. Stein, *Zh.Neorg.Khim.* 6 (1961), 2711–2723.
- [10] J. Mindel, *J. Chem. Edu.* 39 (1962), 512.
- [11] D.R. Rouvray, *Chem. Br.* 10 (1974), 11.
- [12] T.P. Radhakrishnan, *J. Math. Chem.* 5 (1990), 381–387.
- [13] N. Trinajstić, *Chemical Graph Theory* (CRC Press, Boca Raton, 1992), pp. 14–18.
- [14] J. Turulski and J. Niedzielski, *J. Chem. Inf. Comput. Sci.* 42 (2002), 534–539.
- [15] D.Shechtman, I.Blech, D.Gratias, J.W.Cahn, *Phys. Rev. Lett.* 53 (1984), 1951.
- [16] M. Senechal, *Quasicrystal and Geometry* (Cambridge University Press, Cambridge, 1995).
- [17] D. Shechtman and C.I. Lang, *MRS Bulletin*, 22 (1997), 40.
- [18] A.P. Tsai, *MRS Bulletin* 22 (1997), 43.
- [19] J.W. Knop, W.R. Müller, K. Szymanski and N. Trinajstić, *Computer Generation of Certain Classes of Molecules* (Zagreb, SKTH/Kemia, 1985), pp. 28–34.
- [20] F. Harary and E.M. Palmer, *Graphical Enumeration*, Chapter Asymptotics (Academic Press, New York, 1973).
- [21] R.G. Busacker and T.L. Saaty, *Finite Graphs and Networks* (Mc Graw-Hill, New York, 1974) pp. 127–133.
- [22] S.R. Brinkley, *J. Chem. Phys.* 15 (1947) 107–110.
- [23] M. Zhao, Z. Wang and L. Xiao, *J. Chem. Edu.* 69 (1992) 539–542.
- [24] J. Emsley, *Oxford Chemistry Guide: The Elements* (Clarendon Press, Oxford, 1991).