Chemistry-Mathematics-Philosophy Brew: 
a Personal Approach
Alexandru T. BALABAN

Texas A&M University Galveston, 200 Seawolf Park, Galveston, TX 77551

(Received: January 16, 2017; Accepted for publication: March 1, 2017; Online publication: June 6, 2017)

Chemical applications of discrete mathematics and graph theory are briefly reviewed, including philosophical implications. Using the concept of dualist (inner graph-theoretical duals) it was possible to classify (cata-peri-corona classes) and enumerate benzenoid and diamondoid hydrocarbons. By associating numbers with molecular graphs, one can use these numbers (topological indices) for correlations with properties of chemical compounds — an early, simple, and rapid approach to drug design. The two types of atoms (metals and non-metals), are connected by three types pf chemical bonds (ionic, metallic, and covalent) that lead to four types pf lattices (ionic, metallic, atomic and molecular), allowing quick "2-3-4 grasp" of chemistry. The Periodic System of Elements, which is the cornerstone of chemistry and atomic physics, is in danger of being presented wrongly, devoid of the symmetry based on electronic s, p, d, and f shells; several possibilities for showing correctly these shells are discussed.

Keywords: Chemical graphs; Dualists of benzenoids and diamondoids; Classification, codification, and enumeration of benzenoids and diamondoids; Topological indices; 2-3-4 Types of atoms-bonds-lattices; Periodic System of Elements; Electronic s, p, d, and f shells

Immanuel Kant, Auguste Comte, and other philosophers had a low opinion of chemistry because it did not use mathematics. However, Silvester and Cayley were inspired by chemistry which became a third origin of graph theory, along with Euler's topology and Kirchhoff's nets. In turn, graph theory provided a very useful tool for chemical applications because chemistry operates not only with continuous functions (see Dewar's semiempirical approach, Pople's ab initio and Kohn's density functional theory), but also with discrete structures. Polya's theorems provided an elegant solution to finding all isomers for a given molecular formula, and allowing the application of the Sherlock Holmes Principle: [1]: "after you have eliminated the impossible, whatever remains, no matter how improbable, must contain the truth."

Chemists are nowadays familiar with molecular graphs, in which vertices (points) symbolize atoms (usually omitting hydrogens) and edges (lines) symbolize covalent bonds (single, double, or triple). They also make use of the Chemical Abstracts Database containing about $10^8$ structures in order to obtain information about any existing or virtual structure via SciFinder. An early paper with Frank Harary drew attention to the fact that for structural information, the characteristic polynomial of the hydrogen-depleted graph is insufficient [2].

In the first book about graph theory in chemistry [3] and about annulene valence isomers [4], as well many papers in the series "Chemical Graphs" Parts 1 to 53 [5], ample use of Polya's Main Theorem was made, starting with a systematic approach to atoms constituting monocyclic aromatic rings. Depending on the numbers of electrons in the unshared p-orbital of $sp^2$-hybridized atoms, X, Y, and Z type atoms are possible, with 0, 1, or 2 p-electrons. Thus all possible aromatic rings can be found as solutions of the "necklace problem" with three kinds of beads. Similarly, all valence isomers of $[2k]$-annulenes can be found as cubic graphs with $2k$ vertices. Isomers share molecular formulas (e.g., there are over 200 isomers for $C_9H_8$), but valence isomers share partition formulas according to numbers.
of n-ary carbon atoms, e. g. C(CH)4(CH2) or C2(CH)2(CH2)2. For (CH)6 there are only 5 planar cubic graphs.

In addition to molecular graphs, reaction graphs (with vertices symbolizing intermediate molecules and edges symbolizing reaction steps) were introduced in 1966 [6]. The reaction graph for ethyl cations with five different substituents and isotopic labeling is the famous Petersen graph with 10 vertices, or the trivalent 5-cage with an impressive high symmetry. Trivalent g-cages are regular graphs of degree 3 with the minimal circuit (girth) of length 7. As a side-remark, in graph theory 'trivalent' or 'cubic' or 'of degree 3' are synonyms.

All trivalent g-cages with g < 13 are known, but nobody knows till now the cages with g > 12. Also, it is not known why there are 18 nine-cages and 3 ten-cages. The first 10-cage, and the 11-cage, are known as Balaban cages [7].

In 1975, with Professors Oskar E. Polansky, Adalbert Kerber, and André Dreiding, we founded the journal MATCH Communications of Mathematical and Computer Chemistry. It was printed initially in at Polansky's Institute in Mühlheim an der Ruhr, then in Kerber's Institute in Aachen, and is now edited by Professor Ivan Gutman and Dr. Bori Furtula in Kragujevac, Serbia. It concentrates on interdisciplinary research involving chemistry and discrete mathematics.

Another side remark is that in 1975, I had "converted" the young quantum-chemist post-doctoral fellow Milan Randić (who was in the audience at Harvard's Chemistry Department at my lecture), to become a graph-theoretical chemist. With his solid background in mathematics and physics, he made remarkable contributions to mathematical chemistry, and became a life-long friend. Among his many contributions to QSAR he developed a much-cited TI (γ) which was generalized by Lemont B. Kier and Larry H. Hall.

The chemical structure is expressed by constitutional graphs and by stereochemical formulas (diastereo-isomers and enantiomers). The latter type of isomerism leads to very subtle differences in properties involving interaction with chiral forms of energy (e. g. polarized electromagnetic waves) or matter (e. g. chiral solvents or bioreceptors). For drug design, which is the driving force in inventing new medications, one has to devise methods for quantitative structure-property and structure-activity relationships (QSPR and QAR, respectively). One efficient way is to devise topological indices (TIs). This term was introduced by Haruo Hosoya [8] who invited Milan Randić and myself to Japan; in my Tokyo lecture I mentioned that Professor Hosoya is the godfather for TIs. In our culture, godfathers are the men who give children their first names in the Christian ceremony of baptism. However, Japanese people associate the term "godfather" with the movie about Mafia.

Topological indices have several limitations: (i) they are degenerate, i. e. they do not describe uniquely the chemical constitution': (ii) they in most cases increase uniformly with increasing size and cyclicity of the molecule; (iii) they do not encode the diversity of heteroatoms and bond multiplicity unless extra information is added. In order to remedy such limitations, many different approaches have been proposed, one of which resulted in what is known nowadays as the Balaban index J . Among other TIs, the triplet indices that are available through Subhash Basak's computer programs, and the information-based indices U, V, X, and Y may be mentioned. An interesting simultaneous and independent invention may be mentioned: Nenad Trinajstić and his co-workers in Zagreb and our group in Bucharest published what is known as "Harary index" by mutual agreement. These and other TIs are described in a book [9].

Probably the first attempt to compare the energies of alternatives to the graphite and diamond lattices was a publication in 1968 [10]. Benzene (CH)6 is the hydrocarbon corresponding to the unit cell in graphite and graphene, and adamantane (CH)10(CH)6 to the unit cell of the diamond lattice. For brevity, benzenoid hydrocarbons and diamondoid hydrocarbons will henceforth be named benzenoids and diamondoids, respectively. The dual of the sp2-hybridized graphene is the triangulated lattice, whereas the diamond lattice is self-dual. Dualists are not proper graphs because their bond lengths and bond angles do matter, unlike actual graphs. Together with Harary [11] we used the idea of dualist graphs for encoding information on benzenoids: dualists consist in vertices placed in the center of each unit (benzenoid ring) and in edges connecting such vertices in units sharing edges. This idea leads to distinguishing catafusenes, perifusenes and corona-fusenes, according to whether the dualist is acyclic, has 3-membered rings, or respectively larger rings that are not peripherals of 3-membered ring aggregates. On the basis of dualists, it is easy to devise a simple naming system for catafusenes according to whether the dualist increases stepwise from one extreme to the other with angles of
120°, 180°, or 240° denoted by digits 1, 0, and 2, respectively and to a few other conventions, including brackets for denoting branching of the dualist. Benzenoids that have dualists differing only by interconversion of digits 1 and 2 have very similar properties and are called isoarithmic. Although smaller benzenoids have planar molecules, helical benzenoids expanding into the third dimension, called helicenes, are possible. Therefore, enumerating all possible benzenoids have to specify if they do or do not include helicenes.

Stable polycyclic benzenoids are kekuléan, and the stability of isomeric benzenoids may be correlated with the number of Kekulé structures. Major developments in the chemistry of polycyclic aromatic hydrocarbons are due to Eric Clar who prepared over a hundred new benzenoids and who introduced "Clar structures" using the Armit-Robinson circle for π-electron sextets [12]. Unfortunately, there are still authors who ignore the interdiction of writing Clar setets in adjacent rings. Since the publication by Polansky and Derflinger chemist have attempted to assign "degrees of aromaticity" to rings of polycyclic benzenoids. In cases where for one benzenoid there is a Fries structure (a resonance structure with the highest number of aromatic sextets), Clar formulas are equivalent to "winner takes all." For example, in phenanthrene the 14 π-electrons according to its Clar structure are divided among the three rings as 6–2–6. However, by postulating that in averaging for all resonance structures a shared double bond should be divided equally between two rings, one arrives at a partition 5.2–3.6–5.2 whereas the Fries structure corresponds to partition 5–4–5: [13], Figure 1. With D. J. Klein [14], we named "claromatic" the sextet-resonant benzenoids such as triphenylene that have only "full" and "empty" rings as defined by Clar. Dualists helped in finding a method for obtaining all "claromatic" structures, by means of the "anti-sextet duasts" [15]. With the advent of fullerenes and graphene this field exploded in importance. Trying to combine graphene and diamond networks, in 1987 I published a paper with K. M. Merz and Roald Hoffmann about networks that included 3- and 4-connected carbon atoms [16]. A different approach was pursued in collaboration with Douglas J. Klein by attaching pieces of graphite and diamond [17], or exploring nanocones [18].

Dualists also allow a systematical approach to diamondoids. In this case no analogy to helicenes is possible because there

Figure 1. Three views of the distribution of the 14 π-electrons in phenanthrene: A, according to the Fries structure or the Clar sextet rings; B, averaging among all resonance structures; C, assigning 1 and 2 π-electrons, respectively; for shared and unshared double bonds.

Figure 2. Inscribed dualists for benzenoid and diamondoid hydrogen-depleted hydrocarbons. Left: the two isomeric cata-condensed C14H10 anthracene [0], and phenanthrene [1]. Right, the regular [1212]penta-catamantane C26H32 and the irregular [1231]penta-catamantane C25H30.
is no fourth dimension in our real world. Paul Schleyer had found a simple way to synthesize the lower dimondoids sharing carbon hexagons that have only one isomer (adamantane, diamantane, and triamantane) by Lewis-acid-catalyzed isomerization of isomeric hydrocarbons with the same partition. Owing to their lowest steric strain, diamondoids are the most stable among isomers with the same partition. I contacted Paul Schleyer and we published a systematic approach to diamon-doids based on dualists, using the digits 1 to 4 for indicating the four tetrahedral directions around dualist vertices [19]. Similarly, to benzenoids, we can have catamantanes, perimantanes and coronamantanes, but catamantanes can be regular (molecular formula C_{4k+6}H_{4k+10}) and irregular if they have lower numbers of hydrogens. For catamantanes the simple dualist-based code has now been generally accepted, and is being used since the discovery of a natural source of diamonoids in petroleum. Based on dualists, correlations were found between structure and retention times for gas-liquid chromatography and high-performance liquid chromatography [20]. In another collaboration with Douglas J. Klein and his co-workers, we studied theoretically the total energies and the strain energies of acyclic and cyclic aggregates of diamantane units that share a carbon atom, a C–C bond, or a chair-shaped hexagon of carbon atoms [21]. And now to finish with two philosophical-chemical remarks, there is no reason why symmetry should no less important to chemistry than it is to physics and mathematics. For this reason, I believe that the Mendeleev Periodic System of Elements should have a more symmetrical form than the "wallpaper" version adopted by some chemical departments on their walls. This non-symmetrical version breaks the Table after the first three columns starting with H, Be, and Sr, respectively, in order to continue with the f-block, and the remaining columns of the d-block, p-block and s-block. Symmetry-respecting Periodic Systems are presented in Figure 3 in abbreviated form (only with atomic numbers Z), but they should actually contain also the names and symbols of each element. One should read them row after row, from top to bottom, in the natural order of increasing atomic numbers. Although helium would no longer be above the noble gases, the periods (marked by thick lines) should be placed as in the classical form so that each period would end af-

| s  | p  | d  | f  | Z   |
|----|----|----|----|-----|
| 1s | 1,2|     |     |     |
| 2s | 3.4|     |     |     |
| 2p | 5...10|     |     |     |
| 3s | 11,12|     |     |     |
| 3p | 13...18|     |     |     |
| 4s | 19,20|     |     |     |
| 3d | 21...30|     |     |     |
| 4p | 31...36|     |     |     |
| 5s | 37,38|     |     |     |
| 4d | 39...48|     |     |     |
| 5p | 49...54|     |     |     |
| 6s | 55,56|     |     |     |
| 4f | 57...70|     |     |     |
| 5d | 71...80|     |     |     |
| 6p | 81...86|     |     |     |
| 7s | 87,88|     |     |     |
| 5f | 89...102|     |     |     |
| 6d | 103...112|     |     |     |
| 7p | 113,...118|     |     |     |

| f  | d  | s  | p  | Z   |
|----|----|----|----|-----|
| 1s | 1,2|     |     |     |
| 2s | 3.4|     |     |     |
| 2p | 5...10|     |     |     |
| 3s | 11,12|     |     |     |
| 3p | 13...18|     |     |     |
| 4s | 19,20|     |     |     |
| 3d | 21...30|     |     |     |
| 4p | 31...36|     |     |     |
| 5s | 37,38|     |     |     |
| 4d | 39...48|     |     |     |
| 5p | 49...54|     |     |     |
| 6s | 55,56|     |     |     |
| 4f | 57...70|     |     |     |
| 5d | 71...80|     |     |     |
| 6p | 81...86|     |     |     |
| 7s | 87,88|     |     |     |
| 5f | 89...102|     |     |     |
| 6d | 103...112|     |     |     |
| 7p | 113,...118|     |     |     |

Figure 3. Two alternatives for the Periodic System of the elements in short form.
ter the corresponding p-block. In s- and p-blocks (which should be on different rows but always together on adjacent columns), metals and non-metals should be distinguished by different colors, no longer including "semi-metals" as a separate category, so that one has a step-wise demarcation between metals and non-metals. Figure 3 presents two alternatives: the first one is more symmetrical, but the second one is closer to the classical form. A third form with s-f-d-p block sequence was discussed in connection with IUPAC names of elements 113, 115, 117, and 118. By turning top-to-bottom symmetrical forms, one obtains the electronic orbital energy diagram, which results from the fact that electrons, being fermions, obey Pauli's Exclusion Principle.

The second remark makes use of the two notions metals versus non-metals, by a simplification of types of atoms, bonds, and nets; it was published under the title "Chemistry is as simple as 2, 3, 4:. From the two types of atoms we can expect three types of chemical bonds: metallic (metal to metal), ionic (metal cations to non-metal anions) and covalent (non-metal to non-metal). Metallic bonds lead to electrical conductivity in solid and liquid state (similarly to ionic bonds in solutions of electrolytes), whereas covalent bonds characterize insulators or semiconductors. Finally, we have four types of solid nets: metallic and ionic from the corresponding bonds, but covalent bonds lead to the weakest (molecular nets) and strongest (atomic) nets: in the former ones we find molecules (including monatomic noble gas "molecules"), whereas in atomic nets we have tightly-bonded atoms such as carbon in diamond, or silicon and oxygen in quartz.

This simplified version of chemistry is incomplete and better fit for high school, while later for students in universities one can learn subtler details, without the need to "unlearn" false information. In order to avoid wrong guesses, one should replace the expressions "hydrogen or halogen bonds" by "hydrogen or halogen bridges"; hydrogen bridges are strongest intermolecular forces, but they are much weaker than the three chemical bonds [22].

References

[1] A. T. Balaban, Hyle, 19, 106 (2013).
[2] A. T. Balaban, F. Harary, J. Chem. Doc., 11, 258 (1971). DOI:10.1021/c160043a020
[3] A. T. Balaban, (editor), "Chemical Applications of Graph Theory" Academic Press, London, 1976.
[4] A. T. Balaban, M. Banciu, V. Ciorba, "Annulenes, Benzo-, Hetero-, Homo-Derivatives and Their Valence Isomers", CRC Press, Boca Raton, Florida, 1986, 3 volumes.
[5] C. Deleanu and A.T. Balaban, Rev. Roum. Chim., 1990, 35, 349 and previous parts in the series.
[6] A. T. Balaban, D. Farcașiu, R. Banica, Rev. Roum. Chim., 11, 1205 (1966).
[7] A. T. Balaban, Annals Discrete Math., 1993, 55, 109; reprinted in "Quo Vadis, Graph Theory?" (eds. J. Gimbel, J. W. Kennedy and L. V. Quintas), North Holland, Amsterdam, 1993.
[8] H. Hosoya, Bull. Chem. Soc. Jpn., 44, 2332 (1971). DOI:10.1246/bcsj.44.2332
[9] J. Devillers, A. T. Balaban, (eds.), "Topological Indices and Related Descriptors in QSAR and QSPR", Gordon and Breach Sci. Publ., Amsterdam, 1999.
[10] A. T. Balaban, C. C. Rentea, E. Ciupit, Rev. Roum. Chim., 13, 231 (1968).
[11] A. T. Balaban, F. Harary, Tetrahedron, 24, 2505 (1968). DOI:10.1016/S0040-4020(01)82523-0
[12] E. Clar, "The Aromatic Sextet", Wiley, London, 1972.
[13] M. Randić, A.T. Balaban, Polycyclic Arom. Comp., 24, 173 (2004).
[14] A. T. Balaban, D. J. Klein, J. Phys. Chem. C, 113, 19123 (2009). DOI:10.1021/jp9082618
[15] A. T. Balaban, T. G. Schmalz, J. Chem. Inf. Model., 46, 1563 (2006). PMID:16859288, DOI:10.1021/ci0600071
[16] K. M. Merz, Jr., R. Hoffmann, A. T. Balaban, J. Am. Chem. Soc., 109, 6742 (1987). DOI:10.1021/ja00256a031
[17] A. T. Balaban, D. J. Klein, C. A. Folden, Chem. Phys. Lett., 217, 266 (1994). DOI:10.1016/0009-2614(93)E1379-U
[18] D. J. Klein, A. T. Balaban, J. Chem. Inf. Model., 46, 307 (2006). PMID:16426066, DOI:10.1021/ci0503356
[19] A.T. Balaban and P. von R. Schleyer, Tetrahedron, 1978, 34, 3599. DOI:10.1016/0040-4020(78)88437-3
[20] A. T. Balaban, D. J. Klein, J. E. Dahl, R. M. K. Carlson, The Open Org. Chem. J., 1, 13 (2007).
[21] A. T. Balaban, D. Bhattacharya, D. J. Klein, Y. P. Ortiz, Int. J. Quantum Chem., 116, 113 (2016). DOI:10.1002/qua.25031
[22] A. T. Balaban, Chem. Educ., 9, 356 (2004).