

2

Enumeration in Chemistry

BY D. J. KLEIN, D. BABIĆ AND N. TRINAJSTIĆ

1 Introduction and Historical Overview

Overall one might break down the applications of enumeration in chemistry into two broad areas, first an early historically important one concerning isomer enumeration, and second a variety of other enumeration problems arising in detailed descriptions of individual substances. Thence in this section we briefly recall these two broad categories of enumeration, and thereafter attempt a broad view of chemical enumeration, and its future. Thereafter in the ensuing sections we focus more tightly on methodology and recent results.

1.1 Early History: Isomer Enumeration. – The classic type of chemical enumeration concerns the enumeration of possible molecular structures. Indeed the subject predates the understanding of molecular structures, with Humboldt¹ (in 1799) enunciating the possibility of different chemical substances with the same elemental composition. Wöhler² (in 1828) made history converting an inorganic compound (ammonium cyanate) to an isomeric organic compound (urea), thereby setting back a then popular *vitalistic* philosophy. Then later (in 1835) Berzelius³ coined the term *isomer* to describe the circumstance of multiple substances with the same elemental composition. With the proposal of (essentially modern) constitutional formulae (in 1864 by Crum Brown⁴) it was simultaneously noted that different structural formulae associate to different chemical substances. Indeed this was one key point in the acceptance of structural formulae (independent of geometrical embedding). For a decade various structural formulae were generated in an individual manner, without any announced mathematical systematics to the generation of the structures, such as perhaps is quite reasonable for the simpler molecules. Then in 1874 Cayley published his seminal paper⁵ concerning the enumeration of alkanes, with Flavitsky⁶ having made (earlier) independent listings. Discrepancies persisted in the enumeration, in part because of differences in understanding of just exactly what structural formulae might be. But following Cayley some mild activity continued, with finally Henze and Blair⁷ (around 1932) taking the subject of enumeration quite seriously, with tabulations for alkanes and related derivatives for up to around two dozen carbon atoms, with their results appearing in a half-dozen papers. Seemingly astoundingly large

isomer counts of up to over 10^{25} arose, though again errors persisted, in this case because of the tediousness of the involved computations, then done by hand. Then in 1937 Polya wrote his seminal paper⁸ (of >100 pages) concerning enumeration under group action. That is, Polya considered the enumeration of combinatorial mappings from one finite set to another, with there being equivalences amongst the mappings as mediated by group actions on the range set and (more especially) on the image set. In application to the molecular structure problem, for instance, the mappings could be that of Cl atoms and H atoms onto the six ligand positions of hexagonal benzene, with all Cl atoms being equivalent (under any permutation), all H atoms being equivalent (under any permutation), and the six ligand positions being associated to a permutation group corresponding to the rotation and reflections of the hexagonal-symmetry ring of benzene. Polya's theory was described in a quite general mathematical manner, and made a tremendous impact, with his theory now being a standard staple of mathematical combinatorics texts, which often may not mention any chemical examples, though Polya's motivation for the theory was for the purpose of isomer enumeration. Indeed Polya considered the enumeration of alkanes and also published a short illustrative paper⁹ for consumption by chemists. As a side note it is amusing to note that a substantial portion of Polya's formal mathematical results were anticipated in 1927 by J. Redfield. However, Redfield's work was long overlooked,¹⁰ presumably because Redfield used an unusual notation (including astrological symbols) and did not have neat examples of application of the theory. Indeed in 1940 when Redfield submitted a second paper, the journal did not realize any connection with other work and judged the area to be so uninteresting that the paper was rejected. Only after Polya's theory had become established and applied by a number of mathematicians was Redfield's first paper noted (around 1960), and even later was the rejected manuscript found, and finally published¹¹ in 1984, along with a brief survey of Redfield's life and work. Finally in 2000 a yet further more expository manuscript of Redfield's was published.¹² In any event various elaborations and applications of Polya's theory have continued over the years. A comprehensive mathematical monograph¹³ on *Applied Finite Group Actions* by Kerber has recently appeared, with its focus being not only on enumeration under group action, but also on the generation of representatives of the enumerated combinatoric structures. Moreover, this monograph notably unifies various approaches to the enumeration and generation problems, and indeed makes prominent mention of chemical applications, such as seem to have been a prime motivating factor in Kerber's own research. Beyond the enumeration of isomers in isomer classes, Polya's ideas have been applied to other problems of chemical interest – to enumerate: sub- and super-classes of isomers; chemical rearrangement processes; various types of quantum chemical classes of energy levels; and various interaction diagrams.

1.2 Further Enumerations. – Another, somewhat separate area of chemical enumeration concerns the counting of resonance structures for the purpose of gauging the extent of 'resonance'. Often these structures may have much the flavor of different isomeric structures, though the enumeration is usually taken to

be relevant without regard to symmetry equivalences, thereby making some of the underlying mathematical aspects simpler than for the case of isomer enumeration. Again in the earlier work the enumeration was imagined to be done by hand. Pauling's master-work¹⁴ on *The Nature of the Chemical Bond* was written with this view. Especially for the circumstance of conjugated π -electron networks, the numbers could become quite large, say for benzenoids of a dozen or more rings. Thence a more systematic approach became desirable, with a paper by Gordon and Davison¹⁵ (in 1952) providing an early attempt to enumerate Kekule structures for more extended π -networks. With a burgeoning interest in chemical graph theory in the 1970s and thereafter, a few hundred papers enumerating Kekule structures were written, with partial reviews being found in the books by Cyvin and Gutman.¹⁶⁻¹⁸ The matter of making quantitative estimates of resonance energy for conjugated π -networks also came to be addressed in terms of a related enumeration problem of *conjugated circuits*, following the foundational papers of Herndon^{19,20} and of Randić^{21,22} in the 1970s.

Another very broad area of chemical enumeration may be found in the area of statistical mechanics. Thus in the 1930s Mayer²³ developed an elaborate theory of real gases, with graphs characterizing the interactions (and thence the deviations from *ideality*), and their enumeration being crucial to understanding condensation. Moreover there has now developed a whole area of *lattice statistics* which entails numerous types of graph enumeration problems. Often the primary focus is on computing associated enumerative generating functions, which themselves are identifiable as statistical-mechanical partition functions, and various (logarithmic) derivatives thereof giving different thermodynamic properties. Still occasionally there has been focus on explicit enumerations such as in Fowler and Rushbrooke's²⁴ (1936) enumeration of dimer coverings (as can incidentally be identified with Kekule structures); in Montroll's²⁵ (1950) enumeration of partly self-avoiding random walks; in Wall *et al.*'s²⁶ Monte Carlo estimation of counts (and extents) of self-avoiding walks; or in Uhlenbeck and Ford's (1962) review²⁷ of statistical mechanical graph theory and relevant enumeration procedures. Indeed, over the last few decades there have been a few hundred papers dealing with the self-avoiding embedding of linear chains on some regular lattice; a single chain represents a polymer in dilute solution, while more chains represent mixtures, and also other types of graphs to be embedded represent other conceivable (branched) polymer structures (*e.g.* dendrimers). The logarithm of such an enumeration represents a conformational entropy, and again different sorts of weightings and generating functions (and derivatives thereof) are of ultimate interest.

Enumerations also arise in the area of quantum chemistry. Electron-pairing diagrams closely correspondent to chemical structure were early emphasized by Rumer²⁸ (1932), and the computation in terms of these diagrams was soon taken up by Pauling²⁹ (1933) who devised a quantitative scheme entailing the enumeration of certain subgraphs (termed 'islands') obtained by the superposition of electronic pairing diagrams. The resulting valence-bond theory was immediately applied by Pauling and Wheland³⁰ (1933) and thereafter by many others. Here the enumeration problems grow even more rapidly than in simply counting Kekule

structures, in as much as superpositions of pairs of Kekule structures are made, and for each superposition edge and cycle subgraphs are separately enumerated. The ensuing computational difficulty was presumably one of the reasons that valence-bond theory fell into disfavor for a period of time. Now though powerful computational procedures, essentially avoiding the graphical framework, have come to be used in the *ab initio* realm.³¹ And in the semi-empirical realm powerful graph-theoretic techniques have been developed to deal even with quite large systems, such as of relevance in high-temperature superconduction where resonating valence-bond descriptions have been proposed³² (1986) to be of relevance. These graph-theoretic techniques typically focus not on the enumerations but the computation of associated enumerative generating functions, which turn out to give the relevant overall matrix elements. Within the molecular orbital approach, especially with correlation, there are Feynman diagrams³³ and such, which could be enumerated, though typically at issue is their evaluation rather than their enumeration.

There remain yet other chemical enumeration problems. There are various types of knots or links representing knotted or linked molecules, and such may be enumerated, though here too the interest typically goes beyond enumeration (to generation). One may enumerate different types of spectral lines (*e.g.* as expected in a molecular NMR spectrum), though typically more than the number of lines is of interest. That is, in molecular spectroscopies much of the focus typically is on the locations and intensities of the lines and, especially when there are enormous numbers of lines, they may not all be resolved. In the context of various quantitative structure–property relationships (QSPR) and structure–activity relationships (QSAR) there are enumerations of different types of subgraphs which may be used as graph invariants in these correlations. Indeed this area ends up bordering on the vast area of *combinatorial chemistry* and *chemical informatics*, again where the primary focuses typically are beyond enumeration.

1.3 Why Enumerate?. – Clearly enumeration has played an important role in the history of chemistry. But does it still? Are the noted enumerations just historical anachronisms? Is enumeration irrelevant for modern interests in quantitative descriptions of different substances? Indeed in all the areas we have noted, one may indeed argue that enumeration is but a first step towards a more comprehensive characterization and undertaking. Combinatorial formulae often merely identify two different enumerations to have equal values, with one of the enumerations being the easier to perform. We may note for instance that isomer enumeration in Polya theory identifies this enumeration to that of the enumeration of certain equivalence classes of functions. With the counts for two different sets of objects being equal, there often is a natural bijection (*i.e.* a one-to-one correspondence) between the two sets, so that the objects of one set may be used to represent (or even name) those of the other. Thence for the case of chemical isomers again, the mathematical set of objects offers a nomenclature for the isomers. Conversely too, granted a nomenclature, a possibility for enumeration is offered: one seeks to enumerate the names (which presumably exhibit some systematic structure). In some sense then a sensible nomenclature and enumera-

tions can be seen as intimately mutually inter-related. But with systematic labelling (*i.e.* nomenclature) one may anticipate that characterization of the enumerated elements might be systematically made also. In such a case a combinatorial chemistry of (even enormous) virtual libraries of chemical substances might be efficiently handled.

In a number of contexts it seems that transformed enumerations are actually of more central interest. In particular rather than the numbers $\#_N$ of some objects of a size N , interest may focus on an associated generating function

$$F(x) = \sum_N \#_N x^N$$

with x a real variable. In statistical mechanical enumerations such an $F(x)$ may be a partition function with x being a simple function of the ratio of temperature and an interaction strength. For instance, for the Ising model on a lattice network, in a standard combinatorial approach (*e.g.* as explained in ref. 34) the counts $\#_N$ may be identified to the number of (not generally connected) subgraphs of N edges such that every vertex is of even degree, whence $x \equiv \tanh(J/kT)$, with J the (nearest-neighbor) interaction strength, T the absolute temperature, and k Boltzmann's constant. Indeed in such statistical mechanical frameworks, while $F(x)$ might be the partition function of focus, the numbers $\#_N$ might be interpretable as the partition functions in some sort of suitable microcanonical-like ensemble. More generally there can be more than one size characteristic and thence more than one variable. Derivatives of the partition function then lead to various thermodynamic quantities (energies, pressures, specific heats, *etc.*). In closer correspondence with isomer enumeration, the enumeration of conformations is the underlying foundation of *polymer statistics*, with $F(x)$ an associated statistical-mechanical partition function. For suitable enumerations the generating function $F(x)$ may be identified as an overlap for a many-body *cluster-expanded* wave-function, and derivatives of $F(x)$ then giving different matrix elements, including that of the Hamiltonian; see, *e.g.*, ref. 35. In any event enumeration or transformed enumerations underlie quite modern problems.

Generally enumeration may be seen as a beginning wedge into a much wider context of questions and problems. Commonly the motive for enumeration might be efficient means by which to view whole ensembles, the properties of the ensemble, and perhaps properties of selected subcategories. Ultimately enumeration becomes hard to distinguish from combinatorics in general. Thence enumeration has and will remain fundamental to many aspects of chemistry.

2 Enumeration Methods

Methods of enumeration may be sought to be divided into different broad categories. The classical case of enumeration of isomers offers some formal challenge in order to account properly for the different sorts of classifications under various symmetries. Many other types of enumerations offer challenges in

extension to large systems, perhaps infinite (representing fluids or solids). In this section enumeration methodologies are broadly discussed, with some attention to specific examples. Section 3 goes on to describe recent work.

2.1 Enumeration under Group Equivalences. – The enumeration of isomer structures provides the classic example here. Consider a skeleton on which various ligands are to be substituted. Then different substitution patterns may be judged to be equivalent if one may be changed into the other under a symmetry transformation of the skeleton. These transformations may be viewed as permutations on the substitution sites of the skeleton, and include transformations corresponding to skeletal point-group symmetries, and possibly also permutations corresponding to non-rigidities in the molecular skeleton. Thus for an ethane skeleton with six substituent sites, the symmetry group (typically) includes the possibility of pseudo-rotation of the two ends relative to one another. Whether permutations corresponding to improper rotations (reflections, inversions *etc.* – if the skeleton exhibits such) are included or not depends on whether one wishes to distinguish chiral isomers or not. Thence the interest is in enumeration of equivalence classes (each generally containing different numbers of substitution patterns). Most generally there is a symmetry associated to the ligands also: for instance, one can imagine the number of ligand structures to match the number of skeletal substitution sites, but with different subsets of the ligand structures being identical, and so exhibiting a symmetry under interchange of the two. One may imagine that the permutation group for the ligands is naught but a product of symmetric groups $S_{n(i)}$ over the i th set $n(i)$ of n_i equivalent ligands. For example one may be interested in placing one H-atom, two F-atoms and three Cl-atoms onto the ethane skeleton, with a symmetry group taken as $S_{\{1\}} \times S_{\{2,3\}} \times S_{\{4,5,6\}}$. Generally the ligands may themselves exhibit a type of *skeletal* symmetry – for instance for bidentate ligands, they may or may not be symmetric under interchange of the two ends of the ligand, and in interchanging two equivalent bidentate ligands, one needs to simultaneously interchange both ends. For instance, for an octahedral coordination complex with two Cl ligands and two (bidentate) ethylenediamine ligands, the symmetry group of the ligands would be

$$(S_{\{1,2\}} \times S_{\{3,4\}} \times S_{\{5,6\}})\{I, (35)(46)\}$$

(where I is the identity and we use the on-line notation for a permutation, *e.g.* so that $(abc)(de)$ indicates a permutation carrying index $c \rightarrow b$, $b \rightarrow a$, $a \rightarrow c$, $d \rightarrow e$ and $e \rightarrow d$). In general there need not be too much formal mathematical distinction between skeleton and ligand sets – both may consist of disjoint pieces with equivalences between separate pieces, and within any one piece there may be multiple points of attachment with internal symmetries entailing permutation of the attachment point labels. With disjoint pieces there may be chirality changing permutations which rather than changing any one piece into itself instead change one into another, and in addition would act on all pieces (whether skeletal or ligand pieces) simultaneously. One general approach would take the *skeleton* to be the atoms and the *ligands* to be the bonds to be used to

interconnect the atoms. Anyway, given two sets S and L (a skeleton and ligand set) along with symmetries for both, the classical chemical isomer problem is to enumerate the equivalence classes of substitution or addition patterns under the equivalence mediated by the symmetry groups.

The isomer enumeration problem as formulated here is a little different and a little more general than is usually considered. Often (as with Polya^{8,9}) the ligands' internal structure is dismissed; they are presumed to be unidentate, and different types of ligands are identified to different colors. Then instead of correspondences, one may speak in terms of mappings from the skeletal set of sites to the colors, whence a particular value (a color) may be taken several times. The approach indicated here seems to allow a greater generality of view, though not all the formal problems are so neatly solved to date. Even with unstructured ligands there are refinements and decorations to the basic problem so formulated. One may subclassify isomers according to the different symmetries of the substitution pattern; see, *e.g.*, refs. 36 and 37. (Often this substitution-pattern symmetry is that of an intersection between skeletal and ligand symmetries.) One may subclassify isomers according to different substructures in the substituent pattern. For example, in dealing with alkane isomers, a subclassification according to longest subchain (*i.e.* graph diameter) might be considered,³⁸ or subclassification according to number of primary, secondary, tertiary and quaternary carbons might be entertained. In the fundamental isomer enumeration problem with skeleton and ligands respectively being atoms and bonds, 'isomers' consisting of disconnected pieces can arise, so that it is relevant to make a subclassification into connectedness classes (*i.e.* according to the degree of interconnection). Another related type of chemical problem would be to enumerate the number of *rearrangement processes* conceivable for moving ligands around on a given skeleton. In this problem the ligands are viewed as 'passive', so that both the sets S and L are viewed as skeletal structures, and the considered *rearrangement processes* are the equivalence classes of permutations, which are viewed as moving whatever ligand is initially present at a skeletal site to another site. With both skeletons in such a process being the same (connected) structure, the equivalence classes are what has been termed a *polytopal* rearrangement, such as has been considered,^{39,40} or for rearrangements on a trigonal bipyramidal skeleton (or an octahedral skeleton, or other polyhedral skeleton). With the two skeletons different, and potentially disconnected, one faces a rearrangement enumeration problem for general chemical reactions.

Evidently the enumeration problem of equivalence classes under group action is quite general, with Polya's foundational work^{8,9} marking a turning point in the consideration of the problem. Thus perhaps it is reasonable to briefly describe Polya's work, where we have a skeleton with a set S of attachment or substitution sites and a set L of ligands viewed as colors (to be applied to the sites of attachment). The sites of S are acted upon permutatively by the elements of a symmetry group G for the skeleton. The application of the colors (or ligands) of L are viewed to correspond to a mapping f from S into L , with different mappings f and g being equivalent if there is a permutation $\pi \in G$ such that $g = f\pi$ (*i.e.* the result of application of g is equal to the result of application of

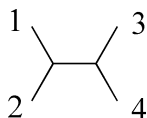
π followed by the application of f to this intermediate result, or also equivalently, g is the functional composition of π with f). Then an equivalence class of these mappings corresponds to an isomer, and the number of different classes are to be enumerated. The numbers of sites of a given color (under such a mapping) correspond to the numbers of different sites occupied by the corresponding ligand, so that we seek equivalence-class counts $\#(\mathbf{n})$ with \mathbf{n} a vector whose α th component n_α gives the number of sites of color α . Polya's solution for this is in terms of a polynomial (called the *cycle index*) associated with the group G ,

$$Z[G; \mathbf{z}] = |G|^{-1} \sum_{\pi \in G} \prod_i z_i^{c_i(\pi)} \quad (1)$$

where $|G|$ is the order of the group, and $c_i(\pi)$ is the number of cycles of size i in the disjoint cycle decomposition of $\pi \in G$. The variables z_i are expressed in terms of a set of color-correspondent dummy variables t_α thus:

$$z_i = \sum_{\alpha \in L} t_\alpha^i$$

Then Polya's renowned *Haupt Satz* gives the various $\#(\mathbf{n})$ as the coefficients of $\prod_{\alpha \in L} t_\alpha^{n_\alpha}$ in $Z[G; \mathbf{z}]$. For example, for the ethylene skeleton (with the 4 H-atoms marking the substitution positions), as



one has $G = D_2 = \{I, (13)(24), (12)(34), (14)(23)\}$ and $Z[D_2; \mathbf{z}] = 4^{-1}\{z_1^4 + 3z_2^2\}$. Then with three possible ligands (say H, F and Cl) one finds

$$\begin{aligned} Z[D_2; \mathbf{z}] &= \{(t_H + t_F + t_{Cl})^4 + 3(t_H^2 + t_F^2 + t_{Cl}^2)^2\}/4 \\ &= t_H^4 + t_F^4 + t_{Cl}^4 + t_H^3 t_F + t_H^3 t_{Cl} + t_F^3 t_H + t_F^3 t_{Cl} + t_{Cl}^3 t_H + t_{Cl}^3 t_F \\ &\quad + 3t_H^2 t_F^2 + 3t_F^2 t_{Cl}^2 + 3t_{Cl}^2 t_H^2 + 3t_H^2 t_F t_{Cl} + 3t_H t_F^2 t_{Cl} + 3t_H t_F t_{Cl}^2 \end{aligned}$$

The isomer counts are in this case readily obtainable even without knowledge of Polya theory, but the method extends to rather much more complicated skeletons. If one were to wish not to distinguish *cis* and *trans* structures, then the group would be extended to $D_2 \cup (12)D_2$.

Beyond this type of case in the preceding paragraph, Polya's work is now much elaborated, reformulated, and extended in numerous fashions, with a recent book¹³ by A. Kerber providing a general mathematical discussion. Though quite comprehensive, Kerber's book may appear mathematically dense to many. On a more chemical side there is the book⁴¹ by Polya and Read, which is a reprinted translation of Polya's (>100-page) original quite lucid article, along with a survey

(by Read) of the literature up through the mid-1980s. The 1991 book by Fujita³⁷ reviews his own elaborated formulation, especially as regards the subclassification with regard to substitution-pattern symmetry. The book⁴² by Trinajstić *et al.* describes the isomer enumeration methodology for substitutional isomers on a fixed skeleton, and it treats especially nicely the case of (acyclic) alkanes, with extensive numerical tabulations. The recent book⁴³ by S. El-Basil surveys the general theory focusing on the circumstance of symmetry subclassification, proceeding in a slow yet quite detailed palatable manner. The overall mathematical area is often referred to as *Polya theory* though much was independently alternatively formulated by Redfield,¹⁰⁻¹² and again there have been other numerous elaborations, reformulations, and extensions.

2.2 Linear Recursive Methods – Kekule Structure Counting. – A rather general class of enumeration problems can be done in a recursive manner, which becomes most especially simplified when the recursion is the same at each stage, corresponding to different local regions of the structure. Thus such approaches are rather readily applicable to the treatment of polymers, crystalline solids, or perhaps other structures of high point-group symmetry. The well-known rotational isomeric model of polymer statistics (*e.g.* as in ref. 44) entails such linear-recursions for a variety of properties, such as mean square radius of gyration. Onsager's famed solution⁴⁵ of the two-dimensional Ising model (on the square lattice) can be viewed as the evaluation *via* a linear recursive technique of a generating function for the enumeration of suitable subgraphs of the lattice – and in this case the exact solution appears quite non-trivial. The enumeration of Kekule structures in conjugated polymers, or in other highly symmetric molecules, such as buckminsterfullerene, provides another example of a linear recursive quantity. More generally there are other possible enumerations, *e.g.* of different subgraph types which may be used as graph invariants in various quantitative structure–property relationships (QSPR) and structure–activity relationships (QSAR). Beyond this, the general subgraph enumeration has other applications, as to statistical mechanics, say as involved with the different Mayer diagrams,²³ or with an enumerative generating function simply being a statistical mechanical partition function. Correlated electronic-structure wave-functions also give rise to a similar generating function problem.³⁵ Again in the bulk of these subgraph enumeration problems equivalence under group action does not play a role and the difficulties dealt with in Polya enumeration theory do not arise, while entirely different problems connected with the subgraph relation arise. Often the subgraph enumeration problem can be solved (at least formally, and perhaps practically) in terms of linear recursions. This broad class of recursive enumeration problems appears to be quite ubiquitous, and in principle soluble by standard linear-algebraic techniques.

As a concrete example, linear recursions have been extensively developed for the case of enumerating Kekule structures. A (molecular) graph might be denoted G , and a subgraph identified as a Kekule structure κ if it has the same number of vertices every one of which has exactly one incident edge in κ . For instance, for naphthalene (Figure 1) one finds three Kekule structures each in correspondence

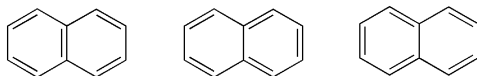


Figure 1 Kekule structures of naphthalene

with one of the possible (neighbor) π -bonding patterns. The greater the number $K(G)$ of such Kekule structures, the greater the resonance stabilization (other things being equal – and in particular for molecular sizes being equal). It seems that the earliest^{15,24} and now quite extensively considered^{16,17} systematic method for Kekule structure enumeration is based on a recursion to smaller graphs. Let e be any edge of graph G , let $G - e$ be the graph obtained from G by deleting e , and let $G \ominus e$ denote the graph obtained by deleting e and all edges incident to e . Then there is a simple recursion

$$K(G) = K(G - e) + K(G \ominus e) \quad (2)$$

which may be readily implemented in a computer program. But also it may be advantageously manipulated in different ways for different special types of graphs. It is noteworthy that if the recursion is iterated with a choice for a sequence of edges so as to disconnect the resultant graphs into disconnected fragments, then the fragments are separately treatable. For polymer graphs it can be used^{15–17,24,46–48} to yield recursions on Kekule-structure counts for chains of different lengths, and perhaps the bulk of Cyvin and co-workers' books^{16–18} is given over to the development of quite explicit formulas as a function of length for a fair number of different particular polymer strips. Even for non-regular polymer graphs and especially catacondensed species there are elegant results, *e.g.* described in Chapter 6 of ref. 17. This latter scheme of Gordon and Davison¹⁵ has a neat pictorial presentation which might be illustrated for a catacondensed polyhex chain:

- first, given the displayed graph, one begins to write in a sequence of numbers in the hexagons starting from one end, with a 2 in the first hexagon and a 1 adjacent to it (in a 0th hexagon);
- second, each number in subsequent hexagons then is the sum of that immediately preceding and the 1st preceding number around a 'kink' in the chain; and
- finally the number appearing in the last hexagon is $K(G)$.

For the example shown in Figure 2, $K(G) = 25$. A natural extension to branched catacondensed species is also known.⁴⁹

There is another neat scheme (based on the linear recursion) applicable for hand computation on polyhex benzenoids of up to a dozen or so hexagonal rings. This *John–Sachs* scheme is based on a one-to-one correspondence between Kekule structures and sets of mutually self-avoiding directed walks on the graph, and indeed this correspondence was (in a special context) utilized⁵⁰ in a statistical mechanical context in modelling collections of partly disordered polymer chains.

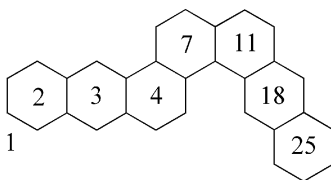


Figure 2 The Gordon–Davison scheme¹⁵ for enumeration of Kekule structures in catacondensed benzenoids

The correspondence between a Kekule structure and a set of three mutually self-avoiding directed walks (from *peak* to *valley*) is illustrated in Figure 3, where the diagonal-oriented double bonds and (bold-face) vertical single bonds are the steps of the walks. Granted this correspondence, John and Sachs⁵¹ developed a neat algorithm to give the Kekule-structure count as the determinant of a small matrix W whose elements count the number of possible directed walks from a position on one side of the polyhex graph G to each position on the other side (independently of any other such walks). Also see ref. 52. For example the path enumerations from either *peak* to either *valley* of the benzenoid structure below are readily obtained as indicated in Figure 4, and then there follows a Kekule-structure count as shown in eqn. (3):

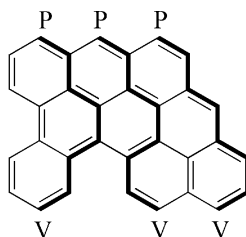


Figure 3 A set of disjoint paths connecting peaks (P) with valleys (V) in a polyhex. The Kekule structure in correspondence with these paths is also indicated

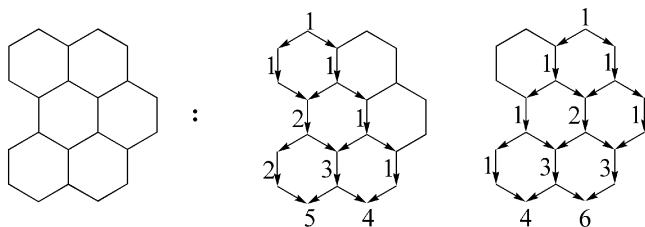


Figure 4 A 'Pascal-triangle' scheme⁵² for enumeration of paths joining the peak with the valleys

$$K(G) = \det(\mathbf{W}) = \det \begin{pmatrix} 5 & 4 \\ 4 & 6 \end{pmatrix} = 14 \quad (3)$$

The consideration of the sets of mutually self-avoiding walkers has a wider impact in revealing a fundamental invariant (the number of walkers) which turns out to be important in transfer-matrix solutions, as applied to polymer graphs extending even to the two-dimensional limit^{53,54} (as discussed in the next subsection). Also this invariant, which may also be described as an *order*, has physical implications.^{55,56}

Overall the general recursion of eqn. (2) is applicable beyond the case of Kekule structures here elaborated for illustration. The related so-called conjugated-circuit method^{20,21,57,58} turns out to have quite neat (related) linear recursions.^{46,47} Generally many sub-graph enumeration problems turn out to be of a linear recursive nature.

2.3 Transfer Matrix Methods. – The linear recursions of the preceding subsection can be alternatively framed in an especially elegant form for polymer graphs, of a rather general type, though here we focus on those polymers which are regular, with the same monomer unit repeated. The Kekule-structure count K_L for a polymer chain of length L monomers can^{59,60} quite generally be cast into the form of a trace

$$K_L = \text{tr}(\boldsymbol{\rho} \cdot \mathbf{T}^L) \quad (4)$$

where \mathbf{T} is a *transfer matrix* characteristic of the monomer unit and $\boldsymbol{\rho}$ is a matrix which encodes the character of the boundary conditions, *i.e.* of the polymer chain ends. Basically one may view \mathbf{T} to give the various ways Kekule structures can propagate from one pattern at the boundary of one monomer unit to another pattern at the subsequent boundary of the next monomer unit. For instance, for a polyphene chain, *e.g.* of 14 hexagons in length, as shown in Figure 5, the chain may be divided into monomer cells of the form shown in Figure 6.

There are just two patterns for the placement of double bonds at the boundary

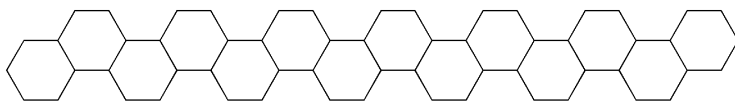


Figure 5 *The polyphene chain*

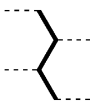


Figure 6 *The monomer repeating in the polyphene chain from Figure 5. Note that the monomer is not only repeated, but also flipped upside-down*

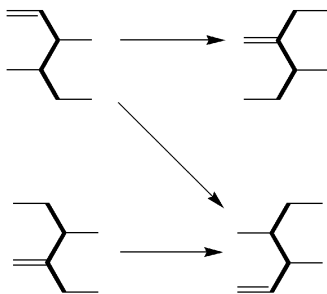


Figure 7 Propagation patterns of the two ways of connecting adjacent monomers in the polyphene chain from Figure 5. The patterns in the left column propagate into those in the right column as indicated by arrows (note that these are flipped upside-down and that the patterns in the left column switched their positions in the right column)

of a unit cell, these two patterns (consistent with types of ends indicated above) being as shown in Figure 7. Of these two patterns, the first may be propagated in two possible manners to the succeeding boundary, while the second may be propagated in but one manner. Thence, for this case, we have the transfer matrix

$$\mathbf{T} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$$

For narrow chains of say a couple hexagons width there are but a few such patterns, so that \mathbf{T} is of a small size (say no more than 4×4), and upon diagonalization of \mathbf{T} one then easily raises \mathbf{T} to the requisite power and a quite neat expression results for arbitrary length L . As strip width w increases (as measured in terms of the number of bonds crossing a monomer boundary), the size of \mathbf{T} increases exponentially with w , but the technique is still readily applicable for widths w up to ~ 12 .

A relation to linear recursions such as of Subsection 2.2 may be indicated. Denote the characteristic polynomial of an $n \times n$ transfer matrix \mathbf{T} by

$$p(x) = \det(x\mathbf{I} - \mathbf{T}) = - \sum_{i=0}^n a_i x^{n-i} \quad (5)$$

Then (by the Hamilton–Cayley theorem) one has

$$\mathbf{T}^n = \sum_{i=1}^n a_i \mathbf{T}^{n-i} \quad (6)$$

whence one can make a substitution (when $L \geq n$)

$$K_{L+1} = \text{tr}(\rho \cdot \mathbf{T}^{L+1-n} \cdot \mathbf{T}^n) = \sum_{i=1}^n a_i \text{tr}(\rho \cdot \mathbf{T}^{L-i+1}) = \sum_{i=1}^n a_i K_{L-i+1} \quad (7)$$

But in fact this is just a linear recursion as in the preceding subsection. For our example polyphene structure one evidently obtains $K_{L+1} = K_L + K_{L-1}$. Notably the recursion is largely independent of the chain ends, which then make their influence with the initial values upon which the recursion is based.

A further means by which to deal with linear recursions and transfer-matrix approaches is by way of generating functions. One introduces generating functions with stage- L counts appearing as coefficients of a dummy variable z raised to the power L . For example, for Kekule structure counting,

$$F(z) = \sum_{L=0} K_L \cdot z^L \quad (8)$$

Then, granted that the transfer matrix satisfies its own characteristic polynomial, one obtains a linear relation for the generating function. For example, for the polyphene chain, where the characteristic polynomial is $x^2 - x - 1$, so that $T^2 = T + I$ (where I is the identity matrix), we obtain

$$\begin{aligned} F(z) &= K_0 + K_1 \cdot z + \sum_{L \geq 2} (K_{L-1} + K_{L-2}) \cdot z^L \\ &= K_0 + K_1 \cdot z + (z + z^2) \cdot F(z) - K_0 \cdot z \end{aligned} \quad (9)$$

Thence one obtains the generating function as a rational polynomial, here

$$F(z) = [K_0 + (K_1 - K_0) \cdot z] / (1 - z - z^2) \quad (10)$$

and this can then be expanded in powers of the dummy variable to obtain explicit expressions for the desired coefficients.

The intimate contact with conjugated circuits theory also may be indicated. This theory concerns a formulation for the resonance energy of molecular structures G which support at least one (fully paired) Kekule structure. Within a Kekule structure κ on G , a conjugated n -circuit is defined to be a cycle (in G) which exhibits alternating pairing in κ around the cycle. Then the resonance energy is

$$E_{\text{res}}(G) = \sum_n \gamma_n \partial_n K(G) / K(G) \quad (11)$$

where the sum is over not too large size n of (even- n) cycles in G , the γ_n are parameters, and $\partial_n K(G)$ is the number of conjugated n -circuits as summed over all Kekule structures κ of G . For regular polymers there is a modification to the transfer-matrix method for K_L so as to also determine the $\partial_n K_L$, and this is well described elsewhere.⁶¹ But also something especially ‘pretty’ happens, granted the

recursion relation for Kekule structure counts of the form in eqn. (7). With the definition of a (squared) polynomial

$$q(x) = [p(x)]^2 = - \sum_{i=0}^{2n} b_i x^{2n-i} \quad (12)$$

and denotation of the conjugated n -circuit count for a length- L chain as $\partial_n K_L$, one finds these conjugated circuit counts satisfy a recursion

$$\partial_n K_{L+1} = \sum_{i=1}^{2n} b_i \partial_n K_{L-i+1} \quad (13)$$

For example polyphene, $q(x) = (x^2 - x - 1)^2$ and

$$\partial_n K_{L+1} = 2 \cdot \partial_n K_L + \partial_n K_{L-1} - 2 \cdot \partial_n K_{L-2} - \partial_n K_{L-3} \quad (14)$$

Notably the coefficients b_i are independent of n (e.g. so that the numerator in the resonance energy expression obeys the same recursion). The distinction of the different conjugated n -circuit counts comes in the initial conditions.

Again all this methodology extends to quite general enumerations. It applies for the Ising model (which may be viewed in essence to be an enumeration of even-degree subgraphs of the parent lattice graph), and it applies to many other statistical mechanical lattice models, as early emphasized by E. Montroll.⁶² Indeed Onsager's famous solution⁴⁵ to the square-planar-lattice Ising model is essentially just a solution of the transfer-matrix eigenproblem, which turns out to be challenging for the extended lattice, because of the dimension of the matrix approaching infinity. In any event the methodology extends to a great variety of graph-theoretic enumerations on polymer graphs.

2.4 Exhaustive Generation (Brute Force) Methods. – Enumeration by explicit generation of all objects being counted is commonly done only when the counts are small or when no better method is (yet) available. The latter situation usually occurs when the underlying structures (which are being counted or on which the counting is being performed) are so irregular that no efficient recursive or group-theoretical methods are applicable. The examples are counting of fullerene isomers,⁶³ fullerene caps,⁶⁴ benzenoid isomers,^{42,65–67} all connected subgraphs of a general graph^{68,69} etc. Usually with such brute-force methods the only shortcut is provided by the *divide-and-conquer* strategy, which however still relies on exhaustive generation on the parts obtained by division. A frequent difficulty in these methods, e.g. when applied to generation of isomers, is an isomorphism of the generated objects, whence this requires additional efforts for diminishing redundancy of the generation algorithm and for recognition of isomorphic objects. In general, though such divide-and-conquer strategies are quite computer intensive, they still have significant advantages over more straightforward approaches. Both sorts of approaches often require computer time increasing

exponentially with system size, though with the divide-and-conquer approach the exponential dependence may be in terms of the subsystem sizes, and thence immensely more efficient at a given size. They typically have a bound on system size which is then only weakly dependent on the computation power available, so that little improvement (in system size) occurs with new generations of computers.

2.5 Other Methods. – An occasional sort of approach is to change one enumeration problem into another. For instance, rather than enumerating Kekule structures, one can consider superpositions of pairs of the structures. These superpositions are still spanning subgraphs, now with components which either are isolated edges or even cycles. The number of such superposition graphs is just the square of the number of Kekule structures, if each cycle in the superposition graphs is identified with a factor of 2 somehow (so that if such a cycle arises from the superposition of Kekule structures κ_a and κ_b , then it in essence also arises from the superposition of κ_b and κ_a , in the reverse order). But in fact such superposition graphs are neatly identified to permutations on G , with the two directions around a cycle rather naturally giving the desired factor of 2 for each cycle. And sums over all the permutations arise in taking the permanent of the adjacency matrix A , the permanent of a matrix being like the determinant, but without the parity of the permutations entering into the sum. Thus

$$\text{per}(\mathbf{A}) = [K(G)]^2 \quad (15)$$

as was noted sometime ago by Percus.⁷⁰ Permanents are generally difficult to compute, whereas determinants are much more convenient, but have a phase problem associated with the parity of the permutations. That is, the determinant of A adds or subtracts superposition graphs, so that one might naturally wonder whether this could be rectified if one were to adjust the signs on the elements of the adjacency matrix. And notably Kasteleyn⁷¹ has found a neat way to solve this phase problem so long as G is planar (in a graph-theoretic sense). This powerful scheme results in a signed adjacency matrix $\mathcal{S}(G)$ with rows and columns that are labelled by the sites of G and with elements that are all 0 except those $\mathcal{S}_{ab} = \pm 1$ with a and b being adjacent sites in G . Then

$$\det(\mathcal{S}(G)) = \pm[K(G)]^2 \quad (16)$$

The signs are such that: first, $\mathcal{S}(G)$ is antisymmetric; second, if for an embedding of G in the plane one follows the edges of \mathcal{S} around any even face (*i.e.* minimal ring) of this embedding, then the number of times a minus sign is encountered is odd. That is, if one proceeds around a ring of sites i_1, i_2, \dots, i_n then

$$\mathcal{S}_{i_1 i_2} \cdot \mathcal{S}_{i_2 i_3} \cdot \dots \cdot \mathcal{S}_{i_{n-1} i_n} \cdot \mathcal{S}_{i_n i_1} = -1 \quad (17)$$

Kasteleyn⁷¹ describes how this *odd orientation* is readily achievable for any planar graph. For instance, if one inserts arrows on edges of G so that an arrow from a

to b indicates $S_{ab} = +1$ while $S_{ba} = -1$, then an example of one such odd orientation is shown in Figure 8.

For the special case of polyhex benzenoid structures, the determinantal formula for $K(G)$ holds except that $A(G)$ appears in place of $S(G)$, as was earlier noted by Dewar and Longuet-Higgins.⁷² Also there is a modification⁷³ of Kasteleyn's scheme by which to deal efficiently with the conjugated-circuit-count problem: one inverts the matrix $S(G)$ and computes determinants for submatrices associated with each cycle for which conjugated-circuit counts are sought. That is, the mean conjugated circuit count for a particular cycle C of a graph G is given as

$$\langle C \rangle_G = \det[(S^{-1})_C] \quad (18)$$

where M_C indicates just the submatrix of M for the rows and columns associated to the sites in C . The method has been applied⁷⁴ to several thousands of fullerenes (many with more than a million Kekule structures), and also⁷⁵ to a couple of dozen different planar carbon network graphs, most with some five- and seven-membered rings.

Really very few significantly different other schemes for computation of Kekule structures or conjugated-circuit counts seem to have been explored to any extent, presumably because the preceding schemes have been so successful. Dewar and Longuet-Higgins⁷² and Herndon⁷⁶ describe an enumeration method based on the identification of non-bonding MOs for radical fragments of the graph under consideration. Randić⁷⁷ has considered a Monte Carlo scheme, but the range of sizes of benzenoids originally intended for this treatment are really quite easily treated exactly by the John-Sachs scheme. Still the Monte Carlo scheme is more evidently extendable to the treatment of the higher-level VB-theoretic models. There are some other less common schemes, *e.g.* as in ref. 78. Various special recursions are discussed in ref. 17.

Moreover, the general idea of mapping one kind of enumeration to another seems to be of fairly general occurrence, though its manifestation typically seems to be quite different in different circumstances. Often there seems to be a correspondence between Ising-model enumerations and Kekule-structure enumerations. For instance, as follows from our brief note in subsection 1.3, for the hexagonal lattice \mathcal{H} the Ising model solution can be viewed as involving enumerations of subgraphs with every vertex having degree 0 or 2. Then one can identify Kekule structures on a corresponding lattice \mathcal{H}^* where each vertex of \mathcal{H}

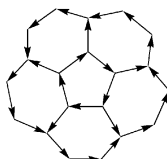


Figure 8 An example of Kasteleyn's⁷¹ odd orientation of a planar polycycle

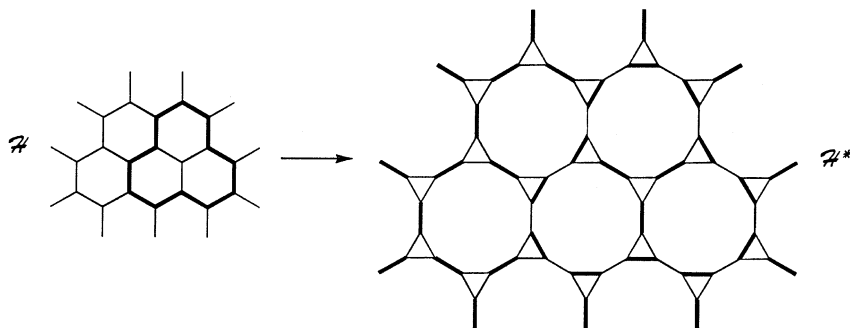


Figure 9 Hexagonal lattice \mathcal{H} and the correspondent lattice \mathcal{H}^* . The subgraph of \mathcal{H} with only even degree vertices corresponds to a unique Kekule structure (perfect matching) in \mathcal{H}^* as indicated in Figure 10



Figure 10 The correspondence between vertices with degrees 2 and 0, in subgraphs of \mathcal{H} , and Kekule structure patterns in \mathcal{H}^*

is replaced by three vertices in \mathcal{H}^* as indicated in Figure 9. Then each Ising-type graph on \mathcal{H} corresponds to a unique Kekule structure on \mathcal{H}^* as is indicated in Figure 10.

Thus Ising-type graph enumeration on \mathcal{H} is equivalent to Kekule-structure enumeration on \mathcal{H}^* , and indeed such correspondences have sometimes been used.⁷⁹ But again correspondences between different types of enumerations seem to be quite frequent. The correspondence of Kekule structures to sets of mutually self-avoiding walks as noted in passing in Section 2.2 in connection with the John–Sachs method of enumeration provides yet another example of such a correspondence.

3 Current Results

Here we survey with some critical commentary recent enumerative work from 1999 through May 2001. We divide up the discussion into different broad categories, separating off the work on fullerene isomer problems from other isomer work, in part because of the activity in dealing with fullerenes, and in part because, to deal best with fullerenes, Polya theory has played a lesser role.

3.1 Isomers: Enumeration and Generation. – During the last two years there seems to have been a degree of activity in the area of isomer characterization, more so with articles devoted to methodology development. There have been

some focusing on isomer enumeration methods. A few of them are classical enumerations for a simple fixed skeleton, with a selection of univalent ligands to be there attached. Lam,⁸⁰ following up on earlier similar work,⁸¹ enumerates the numbers of isomers of alkyl-substituted cyclopropanes – the symmetry is not too high and as a consequence the development may be made without reference to the powerful Polya-theoretic mathematical machinery.

Knopfmacher and Warlimont,⁸² following earlier work,⁸³ use much more esoteric mathematical machinery (concerning semi-groups of polynomial-like structures) to develop methods to determine the asymptotic (large- N) form of isomer counts for N -atom alkanes. Polya⁴¹ in fact developed asymptotic results, but Knopfmacher and Warlimont are interested in extensions, particularly to obtain asymptotics for the numbers of ‘generalized isomers’, viewed as sets of molecules such that different sets as a whole have the same elemental composition. Their main result is a theorem on the asymptotic number of ways in which a mixture of molecules may be realized when the total number of particular kind of atoms is fixed, and the asymptotic results for ordinary single-molecule isomers is already known. Examples of application to mixtures of alkanes, alkenes, substituted alkanes and achiral alkanes and alkanols are given.

Baraldi and Vanossi,⁸⁴ also following up on earlier formal work,⁸⁵ use (in a fairly conventional manner) the general Polya-theoretic machinery to enumerate substitutional isomers for several cyclic or polyhedral skeletons. They conclude with enumerations for icosahedral-symmetry skeletons, both for an icosahedron and for a truncated icosahedron, such as have become of some degree of popularity (as in ref. 86) over the last decade or so because of the ‘elegantly beautiful’ truncated-icosahedral structure of buckminsterfullerene.

J. Szucs⁸⁷ (refining earlier work of Kirby and Pollak⁸⁸) generates elegant near-analytic number-theoretic enumeration formulae for bucky-tori structural isomers (such structures being viewable as finite graphite fragments with cyclic boundary conditions), with special emphasis on asymptotics. Though one can imagine carbon tori with five- and seven-membered rings included (perhaps to contribute⁸⁹ to curvature strain relief), it has been argued⁹⁰ that the experimentally observed⁹¹ (quite large) carbon tori do not have such non-benzenoid rings (at least arranged in any systematic fashion). For molecular polyhedral skeletons, it is well-known that there are naught but one or two topologically reasonable embeddings (with the two arising if the polyhedron is intrinsically chiral). But if for bucky-tori the topological embedding in 3-space is attended to (beyond just the molecular graph), then the same graph may be embedded in different topological manners. For example, a rectangular graphitic sheet may be rolled up in one direction or the other first, as indicated in Figure 11, though either way it is done, the same bonds become connected at the adjoining boundaries (and thence the same graph results). But more than this, even for a fixed order of rolling, one can after the first joining still twist the resultant open tube at the boundary before joining the last boundary, as indicated by the cyclic arrow in the Figure. If the twist is through a multiple of 2π radians, the same graph again results, though the embeddings are generally topologically distinct. Thence the theoretical number of such topologically distinct embeddings grows to infinity (with different multiples

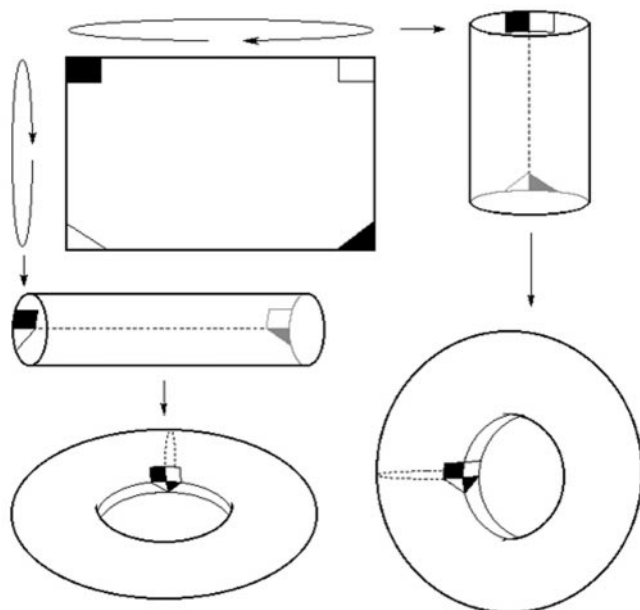


Figure 11 Two different ways of embedding the toroidal graph into a toroidal surface

of 2π), but steric hindrance restricts those which might be plausibly realized, and the labelling of such topological equivalence classes is complete.⁹²

Over the last decade and more S. Fujita has been singularly active in the development of isomer-enumeration methodology and its application. This effort has continued with a variety of further articles. One topic addressed has been^{93–97} that involving isomer enumeration with non-rigid skeletons. Examples are given involving skeletons such as that of tetramethylallene or of dimethylacetylene, with the methyl groups able to undergo internal rotation, while the H-atoms in these methyl groups provide the points of substitution. This involves taking a permutation group expanded over that corresponding to the classical point group, and using this in the Polya-theoretic apparatus. That is, the group is chosen to include permutations corresponding to the internal rotations, or inversions (as at the N-atom of amines) or pseudo-rotations. Perhaps most of Fujita's work over the years has concerned the theory of subsymmetry classification (of the different isomers which arise with different substitution patterns in a skeleton), and he has now further developed^{98,99} this theory, and made applications^{93,100} of it to a number of different circumstances. In yet another sort of circumstance Fujita addresses^{100–102} the enumeration of isomers when the ligands themselves may be chiral, and illustrative applications are made. Here the crucial theoretical point is that permutations (which represent reflections and other improper rotations) need to be recognized as acting on both skeletal positions and the ligands. Further Fujita addresses^{103–106} several interesting chirality characterization questions, concerning 'prochirality', 'stereogenic', 'prostereogenic', 'holotopic', 'hemitopic' *etc.* An-

swers are proposed in terms of Fujita's mathematical constructs, and illustrative applications are made.

S. El-Basil¹⁰⁷ builds from Fujita's work concerning the subclassification of substitutional isomers according to symmetry group of the substitution pattern. He uses so-called *Cayley-like* diagrams as pictorial representations of the actions of group elements on the cosets associated to the various subgroup symmetries, and illustrative applications are made. V.R. Rosenfeld¹⁰⁸ extends Polya theory in two directions: to deal with sets of transformations on skeletal sites such that the set forms just a monoid rather than a full group; and to deal with symmetry subclassification. This latter extension evidently is proposed as an alternative to other approaches, such as Fujita's approach involving *unit induced cycle indices* and *mark tables*, though one can also recognize the mark table in Rosenfeld's *normalizer scheme*. Yet further it may be mentioned that S. El-Basil is to guest-edit a special issue of the journal *MatCh* dedicated to the symmetry subclassification problem for isomer classes. This should appear in 2002.

A few articles¹⁰⁹⁻¹¹¹ have come from a German–Austrian group (of van Almsick, Dolhaine and Hönig) utilizing the classical Polya-theoretic ideas to present substitutional isomer counts for a selection of skeletons, so as to illustrate a general isomer-enumeration program the group has developed. Their article is of an introductory nature insofar as Polya theory is concerned, with the main purpose being to indicate the framework and formatting for their software. Indeed especially one of the co-authors here (H. Dolhaine) has long pursued such a general program, which now is available at:

<http://www-orgc.tu-graz.at/hoegroup>.

Special efficiency in generating libraries of isomer structures is addressed. Another suite of programs due to Kerber's group may be found at:

<http://www.mathe2.uni-bayreuth.de/axel/symneu.engl.html>

and the background for this is discussed in Kerber's monograph.¹³ There are a couple of brief articles¹¹²⁻¹¹⁴ in the chemical literature considering this approach and the software. Both the book and the software are adapted to doing a great variety of different things, besides enumeration also including possibilities for generation, and especially the book gives a wealth of mathematical material, which it seems has only meagerly been utilized to date in chemistry. Further in connection with the journal *MatCh* (from where a number of our references for this subsection come) there is an on-line address:

<http://www.mathe2.uni-bayreuth.de/match/online/links>,

which offers isomer enumeration in a user-friendly format (without necessary reference to formal Polya theory) along with isomer generation (for the first 1000 isomers). The general MOLGEN program to generate isomeric structures, and some related characteristics, is illustratively briefly discussed in ref. 114. A formal

theoretical development of several aspects of this program are discussed in extended detail by Grüner.¹¹⁵ It is available at:

<http://www.mathe2.uni-bayreuth.de/molgen4/>

Hopefully such publicly available resources should find notable use by a wide community of chemists.

Bytautas and Klein have written a set of articles with enumerations of different acyclic hydrocarbon structures: for alkanes;^{116,117} for fully conjugated polyenes;¹¹⁸ and for all hydrocarbons^{119–121} regardless of the degree of unsaturation. Here the alkane enumerations which are included in refs. 116 and 117 repeat or extend earlier work, *e.g.* as reviewed in Trinajstić *et al.*⁴² though in the present articles there are additional considerations, beyond just isomer enumeration. The conjugated polyene enumerations considerably elaborate earlier work¹²² so as now to distinguish radical and non-radical structures – *i.e.* so as to pay attention to the placement of π -bonds. The first article on all acyclic hydrocarbons develops focuses on enumeration, developing the enumeration technology for all acyclic hydrocarbon graphs with arbitrary numbers of double and triple bonds, dealing with the (large-molecule) asymptotics of the isomer counts and applying the methodology for isomers with up to 26 carbons. For this case of all acyclic hydrocarbons, the work is developed with reference to a *formula periodic table*, as indicated in Figure 12.

In this Figure the abscissa is half the number m of H-atoms, while the ordinate is the number n of H-atoms, so that at coordinate (m, n) one finds C_nH_{2m} , and the ordinary alkanes are found on the far right diagonal, the alkenes on the next diagonal in, both alkynes and alkadienes on the third diagonal from the right, *etc.* Moreover, beyond the classical enumerations, these authors seek in the bulk of their work^{116–121} to extend much the same mathematical methodology to compute isomer-class averaged values for different graph invariants, including: atom-type counts, graph diameter, Wiener number, and second moments for atom-type

$\begin{array}{c} m \\ \backslash \\ n \end{array}$	1	2	3	4	5	6
1	• CH ₄					
2	C ₂ H ₂	C ₂ H ₄	C ₂ H ₆			
3	• C ₃ H ₄	C ₃ H ₆	C ₃ H ₈			
4	C ₄ H ₂	C ₄ H ₄	C ₄ H ₆	C ₄ H ₈	C ₄ H ₁₀	
5	• C ₅ H ₄	C ₅ H ₆	C ₅ H ₈	C ₅ H ₁₀	C ₅ H ₁₂	
6	C ₆ H ₂	C ₆ H ₄	C ₆ H ₆	C ₆ H ₈	C ₆ H ₁₀	C ₆ H ₁₂
		⋮				

Figure 12 The formula periodic table of acyclic hydrocarbons

counts (so as to enable determination of associated standard deviations, and cross variances, for atom-type counts). Indeed in ref. 123 results for the Wiener number for alkanes of up to 90 carbons are computed (and asymptotic behaviors considered, not only for all isomers but for other equivalence partitionings of various tree structures). Yet further, the mean graph invariants are utilized along with sub-structural cluster expansions to give estimates for isomer-class averaged values of a few molecular properties, including: heat of formation, magnetic susceptibility, and index of refraction. The techniques thence allow such treatment of isomer classes even with enormous numbers of isomers, *e.g.* $>10^{13}$ for all acyclics¹¹⁹⁻¹²¹ and $>10^{35}$ for just the alkanes.¹¹⁷ For the case of all acyclic hydrocarbons the results are presented¹²⁰ graphically as ‘property overlap plots’, such consisting of contours superimposed on the formula periodic table (of Figure 12). An example of a ‘property overlap plot’ for the heat of formation is found in Figure 13. In this later figure the lines identify constant mean- ΔH_f contours and the arrows normal to the contours indicate magnitudes of standard deviations (for ΔH_f) at these positions, with the scale of the standard deviations set to correspond to the scale associated to the difference between the contour lines. In most of these papers attention is paid to asymptotics, not only for counts but also for the mean values of the various computed graph invariants. One of the articles on alkanes¹¹⁷ also illustrates a rather general method for selecting, from even a very large isomer class, particular structures exhibiting extreme values for the property estimates. It is proposed that the variety of extensions of standard enumerative methodologies considered offers a potential use in screening large classes of structures, in a sort of combinatoric chemistry.

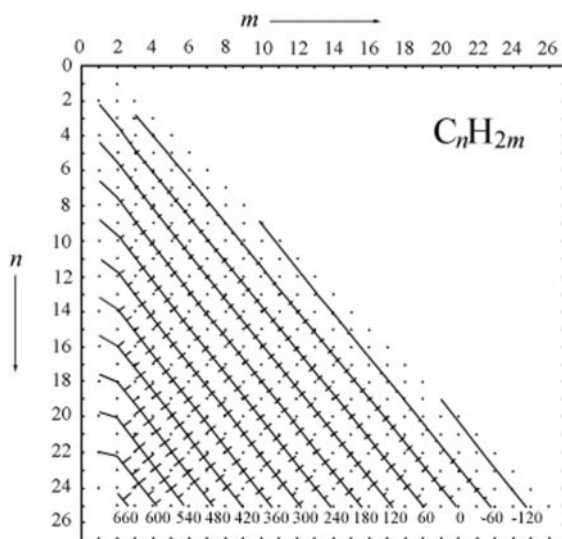


Figure 13 The contour plot of heat of formations for acyclic hydrocarbons C_nH_{2m} . The lines orthogonal to the contours represent standard deviations

There has been some interest in an inverse isomer enumeration problem, of characterizing the symmetry group of the underlying skeleton from the numbers of isomers with different degrees of substitution. Historically this played some role in deducing molecular structures, *e.g.* in seeking a 6-position skeleton for benzene such that there are 1, 3, 3, 3, 1 and 1 chloro-benzenes with respective numbers 1, 2, 3, 4, 5 and 6 of chlorine atoms substituted. Some rather general results were obtained (in 1930) by Lunn and Senior,¹²⁴ later (in 1985) by Hässelbarth¹²⁵ and now by V.V. Iliev.^{126–128} In particular Iliev develops a characterization *via* substitutional reaction processes; with the addition of one more substituent of a given ligand to one isomer with n of these ligands already in place, it may happen that only certain isomers with $n + 1$ of these ligands arise. For instance, chlorination of *p*-dichlorobenzene gives rise to only one (of the three) trichlorobenzenes (while chlorination of *o*-dichlorobenzene gives rise to two trichlorobenzenes, and *m*-dichlorobenzene to all three). At least in favorable cases from the isomer counts (perhaps along with the substitutional reaction diagram just mentioned), the symmetry of the skeleton may be determined. Further Iliev provides explicit example applications for the skeletons of cyclopropane¹²⁸ and of ethane.¹²⁷ However, in the case of benzene it has been emphasized¹²⁹ that, in considering the regular hexagon and the trigonal prism, the problem of determining the skeleton is indeterminate, no matter the number of different (independent monodentate) ligands one considers as substituents, and indeed the problem remains indeterminate even with some additional information about reaction diagrams. That is, granted solely the isomer counts (and reaction-diagram information), both regular hexagonal and trigonal prismatic skeletons for benzene are acceptable.

It may be noted that there are ‘chemical’ problems of different sorts for many of the enumerations which have been made over the years. First, in dealing with the alkanes there is a problem of steric hindrance which typically is entirely ignored. For instance, for alkane-chain conformations, such as treated by Tasi *et al.*^{130,131} and others,¹³² for the longer chains some of the enumerated conformations (thought of as walks on a lattice) end up walking over themselves. Moreover, in the field of ‘polymer statistics’ (concerned not only with conformer enumeration, but also mean spatial extent of these conformers) it is generally accepted (*e.g.* as in ref. 133) that this leads to quantitative differences in predictions as to the mean spatial extent of long-chain polymers (as well as a quantitative difference in the number of allowed conformations). For branched alkanes there are¹³⁴ some conceivable structures for which there exist no self-avoiding embedding on the relevant diamond lattice (even if the spatial requirements of the H-atoms are disregarded), so that one can imagine that the Polya enumeration counts should be reduced. Indeed this fact of steric crowding is¹³⁵ crucial in understanding the experimentally realized termination of dendrimer growth. Another problem arises in the treatment¹¹⁶ of the acyclic polyenes simply as tree structures solely with a limitation on degree (on the H-deleted skeleton to a degree ≤ 3), for then many of the structures turn out to be radicaloid – indeed, as revealed in a ‘corrected’ enumeration¹¹⁰ the great bulk of the otherwise counted *polyenoids* are radicaloid, and even polyradicaloid. Yet further in computing the substitutional

isomers of buckminsterfullerene there is a similar problem, in that many of the enumerated structures do not support a (fully paired) Kekule structure (and thence again are imagined to be radicaloid, and unstable). In some cases of enumeration or generation of isomers there is a ‘reverse’ problem, with some programs counting different Kekule structures of a single isomer as distinct isomers. This might be among the reasons that existing programs for generation of isomers in some cases do not agree in the numbers of produced isomers.¹³⁶ Another circumstance concerns enumerations with bidentate ligands – in this case for many such bidentate species, ligation is restricted to occur at adjacent sites of the skeleton (e.g. ethylenediamine would not ligate to *trans* positions in an octahedral skeleton), while the (extended) Polya-theoretic tools do not account for this. Evidently, even to deal systematically with just enumeration there are a number of ‘chemical’ problems which remain mathematically challenging. In a fair number of articles now many such problems are only overcome in a relatively brute-force explicitly constructional approach.

Contreras *et al.*¹³⁷ have reported an extension of their program CAMGEC,¹³⁸ which generates all isomers from a given molecular formula. The extended version, named US-CAMGEC, is intended to generate the geometrical and stereoisomers induced by cumulative double bonds. Although quite comprehensive, the program does not yet account for chirality of atoms in cycles. The authors proposed an addition to Cahn–Ingold–Prelog rules, but in several comments^{139–142} that followed it was shown to be excessive and contradictory to the existing rules.

In a follow-up of his previous paper,¹⁴³ Le Bret¹³⁶ has reported a comparison of his program *Galvastructures* with other existing programs for generation of isomers. The program *Galvastructures* is unique in using a genetic algorithm for the generation of isomers. Although slower than other programs, especially if all isomers are needed, its advantage is a relatively simple algorithm on which it is based. There were also discussed some practical difficulties in applying the common *fingerpint* methods for recognition of isomorphic structures, and a new fingerprint quantity, based on integer numbers, is proposed.

Lukovits^{144,145} is trying to devise an efficient algorithm for generation of all trees. By using properties of the Morgan labelling,¹⁴⁶ it is easy to generate all the, so-called, Morgan trees. The problem is that to any given isomer corresponds many different Morgan trees, and more so as the number of atoms increases. Thus if one is interested only in isomers, the code becomes highly redundant. However, each isomer has a unique Morgan tree which is used as the canonical code of the isomer and if the generation could be limited to canonical Morgan trees only, the redundancy would be eliminated. Lukovits seeks to formulate a set of semantic rules by which non-canonical Morgan trees could be detected and skipped earlier in the isomer generation. In his recent paper¹⁴⁷ two simple codes of adjacency matrix are formulated and discussed.

An algorithm for generation of all boundary sequences that encompass a planar cubic map consisting of only pentagons or only hexagons was described by Deza *et al.*¹⁴⁸ The algorithm works by producing all sequences satisfying the necessary condition on difference between the numbers of divalent and trivalent boundary vertices. The obtained sequences are subsequently checked for consistency by

deconstruction – that is, by reducing the number of included rings and/or by splitting into smaller sequences which are separately checked in the same way unless the trivial case or some other special case is obtained. An additional, but simpler and faster, algorithm for generating sequences that correspond to planar polyhexes (those embedded into a regular hexagonal lattice) is also described.

Generation and enumeration of benzenoid isomers and fusenes is an everlasting subject considered by many authors. Caporossi *et al.*¹⁴⁹ have counted the numbers of perifused and catafused polyhexes with up to 20 hexagons, and symmetry subcategorization and associated counts are also provided. Chyzak *et al.*⁶⁶ have applied holonomic guessing for estimation of numbers of not yet counted benzenoid isomers: those with 24 and 25 hexagons. The method relies on the built-in procedures of *Maple* and *Mathematica* by which a recurrence relation for a given (presumably holonomic) sequence of numbers can be derived. The predicted number for benzenoids with 24 hexagons agrees with the most recent list of the numbers of benzenoid isomers and fusenes reported by Brinkmann *et al.*⁶⁷ The list contains exact numbers of fusenes with up to 26 hexagons and the number of benzenoids with up to 24 hexagons. Details of the algorithm will be published separately.

An algorithm for generation and enumeration of polycyclic chains was described by Brinkmann *et al.*¹⁵⁰ Polycyclic chain is a graph composed of rings of arbitrary size connected so that its inner dual is a path. Inner dual was also used for condensed representation of the polycyclic chain. A table with 26 representative counts for different combinations of ring sizes and their numbers is also given.

3.2 Fullerenes and Related Objects. – Apart from diversity of their own structural isomers, fullerenes bring additional variety with the derivatives obtained, *e.g.* by reactions of addition. Due to an almost uniform chemical character of all carbon atoms, the ligand atoms could be added in a wide range of numbers and patterns. Possible characteristics of such a reaction have been studied by Fowler *et al.*¹⁵¹ in the model addition of hydrogen onto C₂₄ fullerene. Even with so small a fullerene molecule (the smallest one with hexagonal rings), having also a symmetry, the number of all possible isomers is too large for a systematic and complete study. After abandoning isomers with an odd number of hydrogen atoms, since they imply an open shell in the remaining π -system, the number of isomers (including 0–24 hydrogen atoms) dropped to 352 786. If the isomers containing odd disconnected components in the remaining conjugated π -system (which implies a radical character too) are separated out, the number of isomers to be studied reduces to 63 663 which was considered as acceptable. Nevertheless, due to an efficient method for calculation of electronic structure (density functional based tight-binding method), they were able to calculate energy and perform geometry optimization for all isomers with an even number of hydrogen atoms. The energies of the isomers without explicit radical character (containing no odd component in the conjugated π -system) were distributed at the low end of the cumulative distribution, slightly overlapping with radicaloid species. The most stable isomers for a given number of added hydrogens were

always among those with no odd components. The analysis of the most stable isomers when increasing the number of hydrogens indicated that the hydrogen–carbon bond energy does not change much up to 12 hydrogens, whereafter it decreases. It confirmed Kroto and Walton's¹⁵² suggestion that fullerenes could behave as *superatoms*, exhibiting some sort of preferred coordination that might be interpreted as a preferred valency. In the case of C₂₄, this number turned out to be 12. These twelve hydrogen atoms were bonded along the unique cycle dividing the molecule into two equivalent halves with the bare carbon atoms connected in two hexagonal rings. To check a possible specific effect of hydrogen atom, used as the model ligand, the isostructural most stable isomers with fluorine instead of hydrogen were also evaluated. Again, there was a marked decrease in binding energy at 12 added fluorine atoms showing thus a certain insensitivity to the bonded species. The sequence of the most stable isomers with increasing number of hydrogen atoms is consistent, with no rearrangement of bonded hydrogen atoms, and thus provides a possible pathway of the addition mechanism.

A similar problem has been examined in another paper by Fowler *et al.*¹⁵³ in which there was studied an addition of bromine to the experimentally isolated isomers of C₆₀, C₇₀, C₇₆ and C₈₄. As the bromine atom is much bigger than hydrogen or fluorine atoms, steric interactions between bromine atoms bonded to adjacent carbon atoms make such isomers less stable in comparison to those with no proximal bromines. This allows one to search for the most stable isomer among only those with no adjacent bromine atoms. The number of isomers to be checked in more detail is thus efficiently reduced, especially if only the isomers with maximum numbers of bromine atoms are examined. For example, the number of all isomers of C₆₀Br₂₄ amounts to $\approx 3 \times 10^{14}$, while those with no adjacent bromines there is only 1085. Further reduction is possible if the isomers in which bare carbon atoms do not have a closed-shell electronic structure (as predicted by the Hückel model) are also left out. In this case the number of C₆₀Br₂₄ isomers reduces to a single one. The adducts of C₇₀, C₇₆, and C₈₄ with the maximum number of bromine atoms are determined to be C₇₀Br₂₆ (10 isomers), C₇₆Br₂₈ (36 isomers), C₈₄Br₃₂ (seven isomers and four isomers for two different C₈₄, respectively). Their energies, calculated at the semiempirical level, are also given.

An addition of bulky groups requires more space around each coordinated carbon atom. This may be formalized by a generalized requirement that all added groups must be separated by at least d carbon atoms. Again, for a given molecule and distance d , there may exist many different addition patterns which may not be easy to find. Enumeration and classification of such patterns, named d -codes for a given distance d , has been undertaken in the paper by de la Vaissiere *et al.*¹⁵⁴ for classes of Platonic, Archimedean, face and medial duals of Archimedean polyhedra, as well as for general prism, antiprism and several chemical deltahedra corresponding to carboranes. d -Codes were generated by the program Dense Clique, made by Hansen and Mladenovic.¹⁵⁵ The results are presented in tables containing the numbers and sizes of d -codes for all possible d , classified by symmetry and packing properties.

Face-regular polyhedra, considered in the paper by Brinkmann and Deza,¹⁵⁶ are defined as polyhedra with similar surroundings of equally sized rings. The similar

surrounding means that all n -gonal rings have equal sets of n adjacent rings (with no order implied and rings characterized only by their size), for each n . Written by mathematicians, the paper gives several lists of face-regular polyhedra with constant vertex degrees. The lists are characterized by the size of the maximal ring and the vertex degree, and for some selected cases they are claimed to be complete. Bifaced face-regular polyhedra with constant degree are fully characterized by four theorems. So far the chemical counterparts of regular-faced polyhedra exist in a small number of cases but this may change in the future.

Related icosahedral-symmetry fulleroids are studied by Friedrichs and Deza.¹⁵⁷ Fulleroids have all vertices of degree 3, but sizes of faces other than 5 (and 6) are allowed. Those smaller polyhedral fulleroids with icosahedral symmetry and no more than one size face other than 6 are comprehensively identified. The combinatorial structural characterization of icosahedral fullerenes is addressed by Quinn *et al.*,¹⁵⁸ and shown to be of use in constructions for large such cages. These techniques presumably extend to the icosahedral symmetry fulleroids.

Nanotubes have been theoretically studied by many authors (much of which is reviewed, *e.g.* in ref. 159). However, these studies generally have been concerned only with infinite nanotubes without ends. Brinkmann *et al.*⁶⁴ have studied nanotube caps, that is, the possible ways of ending the tubes, with fullerenic structures. The tube is characterized by the vector (n, m) in the hexagonal lattice connecting two hexagons that overlap each other when rolled into a tube. As there is no unique way to determine a boundary between the tube and the cap, the authors made their own choice by taking the hexagons lying on the two components of the vector (n, m) as the boundary. The patches corresponding to a given boundary were generated by the *divide-and-conquer* method: every patch can be uniquely divided into two special subpatches by using a *Petrie path*. The algorithm constructs all possible special subpatches which are subsequently combined into a patch with the given boundary code. The produced patches are uniquely coded in order to recognize and abandon isomorphic pairs. This is performed for all combinations of n and m , with $n + m \leq 25$, and with $n + m \leq 30$ for patches with isolated pentagons. The patches, whose number exponentially increase with the diameter of the tube, are tabulated by the n and m values and symmetry group. And further the types of asymptotic forms of behaviors for different sorts of ends (regardless of whether the ends are fullerenic or not) are characterized.¹⁶⁰

The papers of Fowler and Rogers¹⁶¹ and of Fowler *et al.*¹⁶² do not deal with enumeration but rather with coding. However, as enumeration is tightly interlaced with coding these papers are also interesting to workers in the field. The first code for fullerenes was devised by Fowler and Manolopoulos *et al.*¹⁶³ It consists of a path going spirally from one face through all other faces. Although not all fullerenes have such a spiral, the smallest known counterexample has 380 vertices¹⁶⁴ and so its use for smaller fullerenes seems to be safe. Particularly for fullerenes with C_5 or C_6 symmetry it has been proven that such spiral exists.¹⁶⁵ In the present paper the authors derived the relation between the Goldberg representation of icosahedral fullerenes and their spiral codes. This relation has been extended also to analogous polyhedra derived from octahedron.

The study of higher genus fullerenoids, particularly with translational symmetry, has been a topic of some interest over the last decade or so. Terrones and Terrones¹⁶⁶ identify some particularly favorable finite structures of genus up to a dozen or so. King¹⁶⁷ also presents some combinatorial group-theoretic aspects of the extended translationally symmetric high-genus such structures.

3.3 Counts of Resonance Structures and Related Items. – Resonance-theoretic based enumerations seem to have been somewhat less studied during the last two years, though a decade or two ago there were tremendous numbers of papers. For the case of Kekule-structure enumeration the methods we deem more powerful or elegant have been briefly indicated in Sections 3.2 and 3.3. Perhaps the current relative quietness of the field indicates that now developed methods are near optimum. Still there has been some work.

Si¹⁶⁸ considers the count of Kekule structures of a special subset of coronoids (which generally are benzenoid structures with a single ‘hole’ region covering two or more hexagons of area). Si establishes a determinantal formula like that of Dewar and Longuet-Higgins⁷² for a special subset of coronoids, and notes that the John–Sachs^{51,52} formula (of Section 2.3) which involves the determinant of a much smaller matrix then also applies for these same coronoids. For more general coronoids there is¹⁶⁹ an extension of the John–Sachs formula, though this entails the evaluation of two or more John–Sachs determinants, and of course the method of Kasteleyn⁷¹ applies.

In a series of papers^{170–174} Dias has derived expressions for resonance-structure counts of several classes of mono- and diradical benzenoid hydrocarbons. The considered classes belong to polymer graphs which were extensively studied some time ago,^{53,59,60,175} though the considered ends of the oligomeric chains may be new. The recursions are apparently obtained by ‘examination’, so that formal derivations of them are not given. The recursions could be more formally achieved by application of the transfer-matrix method for matching polynomials of open polymer graphs (also called *fasciagraphs*),^{176,177} though a slight extension is needed to obtain the first or second derivative of the matching polynomial evaluated at zero. That is, the first derivative of the matching polynomial at zero produces the number of resonance structures for monoradicals, and analogously the k -th derivative (divided by $k!$) gives the resonance-structure count for k -fold polyradicals. For the cases considered the unpaired electrons seem typically to be localized near the chain ends judging from the more numerous valence structures, and some qualitative chemical consequences are considered.

Cash and Herndon¹⁷⁸ described a program for calculation of a matching polynomial. The program is based on a successive removal of edges and application of the well known recursion relation for the matching polynomials. The edges are removed either in a sequence given by the user or automatically – based on the heuristic choice. In order to keep all figures of the matching numbers, the program runs within *Mathematica*.

Zhang and Zhang¹⁷⁹ have strengthened Gutman’s result¹⁸⁰ on the extremal values of Hosoya and Merrifield–Simons indices for hexagonal chains. These two indices are defined as sums of the absolute values of coefficients of the matching

and independence polynomials, respectively. In the present paper it was shown that not only their sum but each coefficient of the two polynomials has extremal values for *linear* and *zigzag* chains. The linear chain has minimal number of matchings and maximal number of independent sets, and *vice versa* holds for the zigzag chain.

The so-called *algebraic structure count* is closely related to Kekule structures but in many cases is more difficult to evaluate. It is defined as the absolute difference of the numbers of two *parities* (*positive* and *negative*) of Kekule structures. This algebraic structure count coincides with $K(G)$ for benzenoids, but for more general alternants it has been advocated as a more reliable measure of aromaticity than $K(G)$. Closed expressions for several special classes have been derived by Gutman *et al.*^{181–184} Recently Graovac *et al.*¹⁸⁵ have formulated a general method for polymers, and elaborated it in detail for a case with two bonds between adjacent monomers. The method uses the fact that the determinant of the adjacency matrix is equal to a square of the algebraic structure count. For a given graph the determinant of the adjacency matrix can be recursively calculated from the determinants of certain subgraphs. Systematic application of such a recursion was efficiently organized by using the transfer matrix technique. The method is illustrated by three classes of phenylenes and two types of acenylenes.

In the paper by Dias¹⁸⁶ electronic structures of several types of subspectral graphs and infinite polymer strips were considered. The common point of enumeration and the subject of this paper are recursions for the characteristic polynomials of one-dimensional polymer graphs. Methods for deriving characteristic polynomial in these cases were first formulated by Polansky and Tyutyulkov¹⁸⁷ by using cyclic symmetry and representative graphs. Later it was formulated in terms of the transfer matrices for polymer graphs with singly connected monomers, and for both open and closed ends as well as for any type of the starting or ending part.¹⁸⁸ Some of these methods seem to be rederived in the present paper.

Lin and Fan¹⁸⁹ have described an algorithm for finding all so-called ‘linearly independent’ and ‘minimal linearly independent’ conjugated circuits in benzenoid hydrocarbons. A different resolution into different-shaped conjugated circuits has also been advocated elsewhere.¹⁹⁰ In another paper¹⁹¹ Lin has used a simpler approach to count only the few smallest conjugated circuits.

3.4 Walks, Connected Subgraphs and Vertices at a Given Distance. – Counts of random walks are used for characterization of graphs and for definition of various molecular descriptors. These counts can be easily obtained from adjacency matrix powers and the recursion provided by the Hamilton–Cayley theorem. Basic mathematical properties of random walk counts were reviewed by Gutman *et al.*¹⁹² In this paper the graphs with extremal walk counts were determined, and the relations between the structure, graph spectra and walk counts were discussed. The spectral moments (that is the numbers of self-returning walks) of phenylenes and their line-graphs were studied by Marković *et al.*¹⁹³ The authors are especially interested in expressing the lower moments by certain

structural details, thus establishing a formal connection between the structure and moments or related molecular descriptors. Formulae for up to the ninth moments of phenylene line-graphs were given in terms of five obvious structural details and compared with similar formulae obtained by Estrada¹⁹⁴ that involved different and less obvious structural counts. A correlation between spectral moments of phenylenes and their line graphs was pointed out.

Rücker and Rücker⁶⁸ have reported an algorithm for the enumeration of all connected subgraphs of a given graph, whose number serves as a molecular descriptor for measuring a complexity of graphs (and molecules represented by them). The algorithm explicitly generates all connected subgraphs by a *depth-first path-tracing* procedure. One might try to devise an algorithm by using a recurrence relation relating the number of connected subgraphs in a graph and in the subgraphs (some of them being ‘rooted’) obtained by removal of an edge. However, such a code would be useless when one wants to know the number of isomorphism classes of connected subgraphs, which is also used for estimation of molecular complexity, but taking into account its symmetry. This was the subject of the next Rücker and Rücker paper,⁶⁹ in which the program for generation of connected subgraphs was combined with calculation of few graph invariants: numbers of vertices and edges, Balaban index, and the extremal eigenvalues of the distance matrix. These invariants were used for (approximate) recognition of isomorphism, so that all subgraphs with the same values of these invariants were considered as isomorphic. The program was generalized for edge- and vertex-colored graphs which are used for representation of heteroatomic molecules and those with multiple bonds. As a spin-off, there were found new examples of graphs with the same Balaban index and of isospectral colored graphs.

The Hosoya polynomial of a graph G is defined as a generating function for the numbers of vertex pairs at varying distance expressed by the exponent of the dummy variable. There is no general recursion to express the Hosoya polynomial of the given graph over its subgraphs, as is possible for *e.g.* matching and characteristic polynomials. Gutman *et al.*¹⁹⁵ have studied several classes of hexagonal chains and formulated (inhomogeneous) recursions for their Hosoya polynomials. These recursions were solved by use of *Mathematica*. Explicit formulas for the Hosoya polynomials were given for members of the considered classes with the number of repeating fragments as a parameter.

3.5 Other Enumerations. – Evaluations of permanents have been pursued by Cash¹⁹⁶ (such permanents being involved in several different enumerations, including that of Kekule structures as mentioned in Section 2.5). He finds efficient means for their evaluation for matrices of up to 80 rows and columns, at least if the matrices have some sparsity.

Tasi *et al.*^{130,131} make an enumeration of conformers of normal-alkane chains. Such does not require use of Polya-theoretic machinery and has been considered several times previously¹³² over the years. A difference in the present enumeration (using relatively straightforward mathematics to give a largely analytic enumera-

tion) is that in addition to the familiar *gauche* and *trans* structures, an evidently theoretically predicted *x*-conformational structure is also included.

Nitta¹⁹⁷ considers some 'polymer statistical' results for the variety of chains of branched acyclic molecules (such as alkanes).

Tissandier *et al.*¹⁹⁸ have studied all possible hydrogen bond arrangements in the water hexamer cage. Although enumeration of hydrogen bond arrangements was recognized as an interesting graph-theoretical problem long ago,^{199,200} only recently²⁰¹ has a simple procedure for generation of all arrangements been outlined. Yet for larger clusters a more efficient program might be necessary which also could analyse the arrangements (as already attempted²⁰²) and the possibilities for their interconversions.

Vismar and Laurenc²⁰³ consider an analysis of a molecular graph into various types of structures. This includes the enumeration and generation of all cycles, in addition to sets of independent cycles. As a further development Dietz *et al.*²⁰⁴ consider discrete decorations of a molecular graph so as to provide additional information about geometric structure, say as regards *cis*- and *trans*-structures, or as regards different enantiomorphs (including different distereomers). The enumeration or generation of the associated discrete mathematical structures then relates to the enumerations or generation of the different chemical isomers, not mediated by ordinary graphs. In fact there are already descriptions of geometric isomerization already implicit in Polya theory, as reviewed in Sections 2.1 and 3.1, though the (permutation-group-theoretical) representations implicit there are generally somewhat different.

Xu and Johnson²⁰⁵ consider the classification of molecules into equivalence classes identified by substructures associated with molecular 'pseudographs'. Here the pseudographs represent homeomorphism classes of graphs (wherein degree-2 vertices are deleted), and again the Morgan extended neighbor idea¹⁴⁶ is utilized.

Further there is a great deal of work involving so-called *topological indices*, which might also be described as molecular graph invariants. Such indices often are integer valued and then count something. For instance the so-called Wiener number of a graph may be viewed as the count of the total number of steps in a set of minimal-length paths one between each pair of distinct vertices of the graph. Various walk counts and the Hosoya index as mentioned in Section 3.4 are further examples of (perhaps less popular) topological indices which enumerate something, and are often so described. The Wiener number however is usually not described as an enumeration, but as the sum over all shortest-path distances. There are numerous other integer-valued topological indices which can be viewed as enumerations, including the Platt index, the Gordon–Scantlebury index, the Zagreb group indices, centric indices, the Szeged index, the hyper-Wiener index, *etc.* And further there are a fair number of topological indices which take rational-number values but for which the numerator and denominator of the rational number both can be viewed as enumerations. There has in fact been a fair degree of work on such topological indices, but we have not attempted to review it fully. It is appropriate to note though that there are recent books^{206–208} which review the field.

4 Conclusion

It is seen that there is a great deal of activity in chemical enumeration. Perhaps in the classical area of isomer enumeration there is even a surprising amount of activity. This may be due to a recognition that it is desirable to treat whole sets of molecules beyond just individual molecules one at a time as has been the dominant focus in quantum chemistry and perhaps even the vast bulk of chemistry over the last several decades. Indeed in general chemistry, and especially medicinal chemistry, this recognition is involved in the intense interest in combinatorial libraries. Much of this area of *combinatorial chemistry* is viewed to be purely concerned with experimental methods of synthesizing large sets (*i.e.* ensembles) of compounds and their ensuing testing for their properties or bio-activities. But then perhaps the current theoretical activity is a quest for some sort of theoretical analogue approaches to comparably deal with virtual (*i.e.* theoretical or computer generated) ensembles of molecules to examine. As witnessed in Section 3.4 there has been much successful work in developing libraries of molecular structures, especially in Brinkmann's work. And the member structures of such libraries may be examined structure by structure, as exemplified in work by Fowler and colleagues. But beyond individual examination of each member, there are other systematic possibilities, which are beginning to be examined and which may ultimately prove quite useful. If a property is (approximately) expressed in terms of sufficiently nice graph invariants (as by *e.g.* cluster expansion in terms of local substructural counts), then searching through graphs, where the graph approximant is near optimal, could be addressed in a way that avoids structure by structure examination.

There also seems still to be much room for theoretical methodological developments. We have noticed (*e.g.* as in Section 2.5) correspondences with seemingly different types of enumerations. And thence there is a suggestion of a general classification of interconnections, with some sort of set of 'canonical' enumerations. Indeed the point of P- or NP-complete algorithms (as for enumerations) made in the mathematical theory of computation presumably focuses on some aspect of such correspondences, and in statistical mechanics the idea of 'universality classes' presumably relates to this also. But the mathematical theory relates to maximum computational time for a general case (*e.g.* enumeration for a particular type of subgraph on a general graph) and the statistical-mechanical ideas relate to asymptotic behaviors, whereas it seems that often there is a much more explicit correspondence between types of enumerations. Perhaps much more of such relations will become apparent in future work.

Overall there seems to have been a degree of activity in enumeration, with indications of a number of further developments to come.

References

1. A. von Humboldt, *Versuche über die gereizte Muskel- und Nervenfasern, nebst Vermutungen über den chemischen Prozess in der Tier- und Pflanzenwelt*, 1797, as quoted in A. Kerber, *Comm. Math. Comput. Chem.*, 1999, **39**, 127.

2. F. Wöhler, *Poggendorf's Ann. Phys.*, 1828, **12**, 253.
3. J. Berzelius, *Ann. Phys. Chem.*, 1830, **19**, 305.
4. C. Brown, *Trans. Roy. Soc. Edinburgh*, 1864, **23**, 707.
5. A. Cayley, *Phil. Mag.*, 1874, **47**, 444; A. Cayley, *Ber. Dtsch. Chem. Ges.*, 1875, **8**, 1056.
6. F. Flavitsky, *J. Russ. Chem. Soc.*, 1871, 160.
7. C.M. Blair and H.R. Henze, *J. Am. Chem. Soc.*, 1932, **54**, 1098, 1538.
8. G. Polya, *Acta Math.*, 1937, **68**, 145.
9. G. Polya, *Zeit. Kryst. A*, 1936, **93**, 414.
10. J.H. Redfield, *Am. J. Math.*, 1927, **49**, 433.
11. J.H. Redfield, *J. Graph Theory*, 1984, **8**, 205 (written in 1940).
12. J.H. Redfield, *Comm. Math. Comput. Chem.*, 2000, **41**, 7 (typescript of a lecture, 1937).
13. A. Kerber, *Applied Finite Group Actions*, Springer-Verlag, New York, 1999.
14. L. Pauling, *The Nature of the Chemical Bond*, Cornell University Press, Ithaca, NY, 1939.
15. M. Gordon and W.H.T. Davison, *J. Chem. Phys.*, 1952, **20**, 428.
16. I. Gutman and S.J. Cyvin, *Introduction to the Theory of Benzenoid Hydrocarbons*, Springer-Verlag, Berlin, 1989.
17. S.J. Cyvin and I. Gutman, *Kekule Structures in Benzenoid Hydrocarbons*, Springer-Verlag, Berlin, 1988.
18. S.J. Cyvin, J. Brunvoll, and B.N. Cyvin, *Theory of Coronoid Hydrocarbons*, Springer-Verlag, Berlin, 1991.
19. W.C. Herndon, *J. Am. Chem. Soc.*, 1973, **95**, 2404.
20. W.C. Herndon, *Thermochim. Acta*, 1974, **8**, 225.
21. M. Randić, *J. Am. Chem. Soc.*, 1977, **99**, 444.
22. M. Randić, *Tetrahedron*, 1977, **33**, 1905.
23. J. Mayer, *J. Chem. Phys.*, 1937, **5**, 67.
24. R.H. Fowler and G.S. Rushbrooke, *Trans. Faraday Soc.*, 1937, **33**, 1272.
25. E. Montroll, *J. Chem. Phys.*, 1950, **18**, 734.
26. F.T. Wall, L.A. Hillier, Jr. and D.J. Wheeler, *J. Chem. Phys.*, 1954, **22**, 1036.
27. G.E. Uhlenbeck and G.W. Ford, in *Studies in Statistical Mechanics*, Vol. I, ed. J. deBoer and G.E. Uhlenbeck, North Holland, Amsterdam, 1962, pp. 119–211.
28. G. Rumer, *Gött. Nach. ges. Wiss.*, 1932, 337.
29. L. Pauling, *J. Chem. Phys.*, 1933, **1**, 280.
30. L. Pauling and G.W. Wheland, *J. Chem. Phys.*, 1933, **1**, 362.
31. E.g. J. Gerratt, D.L. Cooper and M. Raimondi, in *Valence-Bond Theory and Chemical Structure*, ed. D.J. Klein and N. Trinajstić, Elsevier, Amsterdam 1990, pp. 287–350.
32. P.W. Anderson, *Science*, 1987, **287**, 1196.
33. R.P. Feynman, *Phys. Rev.*, 1949, **76**, 749.
34. C.J. Thompson, *Mathematical Statistical Mechanics*, Macmillan Co., New York, 1972.
35. E.g. D.J. Klein, H. Zhu, R. Valenti and M.A. Garcia-Bach, *Int. J. Quantum Chem.*, 1997, **65**, 421.
36. W. Hässelbarth, *Theor. Chim. Acta*, 1985, **67**, 339; J. Brocas, *J. Am. Chem. Soc.*, 1986, **108**, 1135; A. Meade, *J. Am. Chem. Soc.*, 1987, **109**, 2130.
37. S. Fujita, *Symmetry and Combinatorial Enumeration in Chemistry*, Springer-Verlag, Berlin, 1991.
38. A.T. Balaban, J.W. Kennedy and L. Quintas, *J. Chem. Ed.*, 1988, **65**, 304.
39. W.G. Klemperer, *J. Am. Chem. Soc.*, 1972, **94**, 8360; W. Hässelbarth and E. Ruch, *Theor. Chim. Acta*, 1973, **29**, 259.

40. J. Brocas, M. Gielen and R. Willem, *The Permutational Approach to Dynamic Stereochemistry*, McGraw-Hill, New York, 1983.
41. G. Polya and R.C. Read, *Combinatorial Enumeration of Groups, Graphs, and Chemical Compounds*, Springer, New York 1987.
42. N. Trinajstić, S. Nikolić, J. V. Knop, R. W. Müller and K. Szymanski, *Computational Chemical Graph Theory: Characterization, Enumeration and Generation of Chemical Structures by Computer Methods*, Ellis Horwood Ltd., Simon and Schuster, 1991.
43. S. El-Basil, *Computational Organic Chemistry, An Educational Approach*, Nova Science Publishers, 2000.
44. P.J. Flory, *Statistical Mechanics of Chain Molecules*, Interscience, NY, 1969.
45. L. Onsager, *Phys. Rev.*, 1944, **65**, 117.
46. M. Randić, *Int. J. Quantum Chem.*, 1980, **27**, 549.
47. T.P. Živković, M. Randić, D.J. Klein, H. Zhu and N. Trinajstić, *J. Comp. Chem.*, 1995, **16**, 517.
48. J.L. Hock and R.B. McQuistan, *Discrete Appl. Math.*, 1984, **8**, 101; H. Hosoya and A. Motoyama, *J. Math. Phys.*, 1985, **26**, 157; A. J. Phares, *J. Math. Phys.*, 1984, **25**, 1756.
49. D.J. Klein and W.A. Seitz, *J. Mol. Struct. (Theochem)*, 1988, **169**, 167.
50. J.F. Nagle, *Phys. Rev. Lett.*, 1975, **34**, 1150; C. S. O. Yokoi, J. F. Nagle and S. R. Salinas, *J. Stat. Phys.*, 1986, **44**, 729.
51. P. John and H. Sachs, Wegesysteme und Linearfaktoren in hexagonalen und quadratischen Systemen, in: *Graphen in Forschung und Unterricht*, Springer-Verlag 1985, pp. 80–95.
52. P. John and H. Sachs, Calculating the numbers of perfect matchings and of spanning trees, Pauling's orders, the characteristic polynomial, and the eigenvectors of a benzenoid system, in: *Advances in the Theory of Benzenoid Hydrocarbons*, Springer-Verlag, 1990, pp. 145–179.
53. D.J. Klein and T.P. Živković, *Math. Comp. Mod.*, 1993, **17**, 113.
54. H. Sachs, P. Hansen and M. Zheng, *Comm. Math. Comput. Chem.*, 1996, **33**, 169.
55. D.S. Rokhsar and S.A. Kivelson, *Phys. Rev. Lett.*, 1988, **61**, 2107; S. A. Kivelson, *Phys. Rev. B*, 1989, **39**, 259.
56. D.J. Klein, T.G. Schmalz, M.A. Garcia-Bach, R. Valenti and T.P. Živković, *Phys. Rev. B*, 1991, **43**, 719.
57. L.J. Schaad and B.A. Hess, *Pure Appl. Chem.*, 1982, **54**, 1097.
58. D.J. Klein, *J. Chem. Educ.*, 1992, **69**, 691.
59. D.J. Klein, T.G. Schmalz and G.E. Hite, *J. Comp. Chem.*, 1986, **7**, 443.
60. D. Babić and A. Graovac, *Croat. Chem. Acta*, 1986, **59**, 731.
61. D.J. Klein, W.A. Seitz and T.G. Schmalz, pages in *Computational Chemical Graph Theory*, ed. D. Rouvray, Nova Science Publishers, New York, 1990, pp. 128–151.
62. E. Montroll, *J. Chem. Phys.*, 1941, **9**, 706.
63. G. Brinkmann and A.W.M. Dress, *J. Algorith.*, 1997, **23**, 345.
64. G. Brinkmann, P.W. Fowler, D.E. Manolopoulos and A.H.R. Palser, *Chem. Phys. Lett.*, 1999, **315**, 335.
65. W.R. Müller, K. Szymanski, J.V. Knop, S. Nikolić and N. Trinajstić, *J. Comput. Chem.*, 1990, **11**, 223.
66. F. Chyzak, I. Gutman and P. Paule, *Comm. Math. Comput. Chem.*, 1999, **40**, 139.
67. G. Brinkmann, G. Caporossi and P. Hansen, *Comm. Math. Comput. Chem.*, 2001, **43**, 133.
68. G. Rucker and C. Rucker, *Comm. Math. Comput. Chem.*, 2000, **41**, 145.
69. G. Rucker and C. Rucker, *J. Chem. Inf. Comput. Sci.*, 2001, **41**, 314; (Errata: *J. Chem. Inf. Comput. Sci.*, 2001, **41**, 865).

70. J.K. Percus, *J. Math. Phys.* 1969, **10**, 1881.
71. P.W. Kasteleyn, Chapter 2 in *Graph Theory and Theoretical Physics*, ed. F. Harary, Academic Press, New York, 1967.
72. M.J.S. Dewar and H.C. Longuet-Higgins, *Proc. Roy. Soc. A*, 1952, **214**, 482.
73. D.J. Klein and X. Liu, *J. Comp. Chem.*, 1991, **12**, 1260.
74. X. Liu, D.J. Klein and T.G. Schmalz, *Fullerene Sci. Technol.*, 1994, **2**, 405.
75. X.Y. Zhu, A.T. Balaban and D.J. Klein, *J. Chem. Phys.*, 1994, **101**, 5281.
76. W.C. Herndon, *Tetrahedron*, 1973, **29**, 3.
77. M. Randić, *Chem. Phys. Lett.*, 1986, **128**, 193.
78. C. Rongsi, S.J. Cyvin, B.N. Cyvin, J. Brunvoll and D.J. Klein, *Topics Curr. Chem.*, 1990, **153**, 228.
79. M.E. Fisher, *J. Stat. Phys.*, 1984, **34**, 667.
80. C.W. Lam, *J. Math. Chem.*, 2000, **27**, 23.
81. A.T. Balaban, *Croat. Chem. Acta*, 1978, **51**, 35.
82. J. Knopfmacher and R. Warlimont, *J. Math. Chem.*, 1999, **26**, 95.
83. J. Knopfmacher, *J. Math. Chem.*, 1998, **24**, 61.
84. I. Baraldi and D. Vanossi, *J. Chem. Inf. Comput. Sci.*, 2000, **40**, 386.
85. I. Baraldi, C. Fiori and D. Vanossi, *J. Chem. Inf. Comput. Sci.*, 1999, **25**, 23.
86. S. Fujita, *Bull. Chem. Soc. Jpn.*, 1991, **64**, 3215; K. Balasubramanian, *Chem. Phys. Lett.*, 1991, **183** 292; K. Balasubramanian, *J. Phys. Chem.*, 1993, **97**, 6990.
87. J.M. Szucs, *Discrete Math.*, 2000, **215**, 225.
88. E.C. Kirby and P. Pollak, *J. Chem. Inf. Comput. Sci.*, 1998, **38**, 66.
89. S. Itoh, S. Ihara and J. Kitakami, *Phys. Rev. B*, 1993, **47**, 1703, 12908; E.G. Gal'pern, I.V. Stankevich and A. Chistyakov, *Fullerene Sci. Technol.* 1994, **2**, 1; B. Borštnik and D. Lukman, *Chem. Phys. Lett.*, 1994, **228**, 312.
90. D. Babić, D.J. Klein and T.G. Schmalz, *J. Mol. Graphics and Mod.*, 2001, **19**, 222.
91. J. Liu, H.J. Dai, J.H. Hafner, D.T. Colber, R.E. Smalley, S.H. Tans and C. Dekker, *Nature*, 1997, **385**, 780; M. Ahlskog, E. Seynaeve, R.J.M. Vullers, C. van Haesendonck, A. Fonseca, K. Hernadi, and J.B. Nagy, *Chem. Phys. Lett.*, 1999, **300**, 202; R. Martel, H.R. Shea and P. Avouris, *Nature*, 1999, **398**, 299.
92. J. Szucs and D.J. Klein, *Discrete Appl. Math.*, 2000, **105**, 225.
93. S. Fujita, *Bull. Chem. Soc. Jpn.*, 1999, **72**, 1759.
94. S. Fujita, *Bull. Chem. Soc. Jpn.*, 1999, **72**, 2409.
95. S. Fujita, *Bull. Chem. Soc. Jpn.*, 1999, **72**, 2403.
96. S. Fujita, *J. Chem. Inf. Comput. Sci.*, 2000, **40**, 135.
97. S. Fujita, *J. Chem. Inf. Comput. Sci.*, 2000, **40**, 426.
98. S. Fujita, *Theor. Chem. Acc.*, 1999, **101**, 409.
99. S. Fujita, *Bull. Chem. Soc. Jpn.*, 2000, **73**, 329.
100. S. Fujita, *J. Chem. Inf. Comput. Sci.*, 1999, **39**, 151.
101. S. Fujita, *J. Chem. Inf. Comput. Sci.*, 2000, **40**, 1101.
102. S. Fujita, *Bull. Chem. Soc. Jpn.*, 1999, **72**, 13.
103. S. Fujita, *Bull. Chem. Soc. Jpn.*, 2000, **73**, 2679.
104. S. Fujita, *Bull. Chem. Soc. Jpn.*, 2000, **73**, 1979.
105. S. Fujita, *Tetrahedron*, 2000, **56**, 735.
106. S. Fujita, *Bull. Chem. Soc. Jpn.*, 2000, **73**, 2009.
107. S. El-Basil, *J. Chem. Inf. Comput. Sci.*, 2000, **40**, 572.
108. V.R. Rosenfeld, *Comm. Math. Comput. Chem.*, 2001, **43**, 111.
109. H. Dolhaine, H. Höning and M. van Almsick, *Comm. Math. Comput. Chem.*, 1999, **39**, 21.
110. M. van Almsick, H. Dolhaine and H. Höning, *J. Chem. Inf. Comput. Sci.*, 2000, **40**, 956.

111. M. van Almsick, H. Dolhaine and H. Hönic, *Comm. Math. Comput. Sci.*, 2001, **43**, 153.
112. A. Kerber, R. Laue, and T. Wieland, Discrete Mathematics for Combinatorial Chemistry, in *Discrete Mathematical Chemistry*, DIMACS Series in Discrete Mathematical and Theoretical Computer Science, Vol. 51, eds. P. Hansen, P.W. Fowler and M. Zheng, American Mathematical Society, Providence, RI, 2000, pp. 225–234.
113. A. Kerber and A. Kohnwer, *Comm. Math. Comput. Chem.*, 1998, **38**, 163.
114. T. Grüner, A. Kerber, R. Laue and H. Meyer, *Comm. Math. Comput. Chem.*, 1999, **39**, 135.
115. T. Grüner, *Comm. Math. Comput. Chem.*, 1999, **39**, 39.
116. L. Bytautas and D.J. Klein, *J. Chem. Inf. Comput. Sci.*, 1998, **38**, 1063.
117. L. Bytautas and D.J. Klein, *J. Chem. Inf. Comput. Sci.*, 2000, **39**, 803.
118. L. Bytautas and D.J. Klein, *Theor. Chem. Acc.*, 1999, **101**, 371.
119. L. Bytautas and D.J. Klein, *Croat. Chem. Acta*, 2000, **73**, 331.
120. L. Bytautas and D.J. Klein, *Phys. Chem. Chem. Phys.*, 1999, **1**, 5565.
121. L. Bytautas, D.J. Klein and T.G. Schmalz, *New J. Chem.*, 2000, **24**, 329.
122. S.J. Cyvin, J. Brunvoll and B.N. Cyvin, *J. Mol. Struct. (Theochem)*, 1995, **357**, 255; C. Yeh, *J. Chem. Inf. Comput. Sci.*, 1995, **35**, 912; S.J. Cyvin, J. Brunvoll, E. Brendsdal, B.N. Cyvin and E.K. Lloyd, *J. Chem. Inf. Comput. Sci.* 1995, **35**, 743; C. Yeh, *J. Phys. Chem.*, 1996, **100**, 15800.
123. L. Bytautas and D.J. Klein, *J. Chem. Inf. Comput. Sci.*, 2000, **40**, 471.
124. A.C. Lunn and J.K. Senior, *J. Phys. Chem.*, 1929, **33**, 1027.
125. W. Hässelbarth, *J. Comput. Chem.*, 1987, **8**, 700.
126. V.V. Iliev, *Comm. Math. Comput. Chem.*, 1999, **40**, 153.
127. V.V. Iliev, *Comm. Math. Comput. Chem.*, 2001, **43**, 67.
128. V.V. Iliev, *Comm. Math. Comput. Chem.*, 2001, **43**, 79.
129. D.J. Klein and L. Bytautas, *Comm. Math. Comput. Chem.*, 2000, **42**, 261.
130. G. Tasi and F. Mizukami, *J. Math. Chem.*, 1999, **25**, 55.
131. G. Tasi, F. Mizukami, J. Csontos, W. Györfy and I. Palinko, *J. Math. Chem.*, 2000, **27**, 191.
132. E. Funck, *Zeit. Elektrochem.*, 1958, **62**, 901; A.T. Balaban, *Rev. Roum. Chim.*, 1976, **21**, 1049; J. Brunvoll, B.N. Cyvin, E. Brendsdal, and S.J. Cyvin, *Computers Chem.* 1995, **19**, 379; S.J. Cyvin, J. Brunvoll, B.N. Cyvin, and E. Brendsdal, *Adv. Mol. Struct. Res.* 1995, **2**, 213.
133. E.g. P.G. deGennes, *Scaling Concepts in Polymer Physics*, Cornell University Press, Ithaca, New York, 1979.
134. D.J. Klein, *J. Chem. Phys.*, 1981, **75**, 5186.
135. E.g. J.M.J. Frechet, *Science*, 1994, **263**, 1710; D. Tzalis and Y. Tor, *Tetrahedron Lett.*, 1996, **37**, 8293; F. Zeng and S.C. Zimmerman, *Chem. Rev.*, 1997, **97**, 1681.
136. C. Le Bret, *Comm. Math. Comput. Chem.*, 2000, **41**, 79.
137. M.L. Contreras, G.M. Trevisiol, J. Alvarez, G. Arias and R. Rozas, *J. Chem. Inf. Comput. Sci.*, 1999, **39**, 475.
138. M.L. Contreras, R. Rozas and R. Valdivia, *J. Chem. Inf. Comput. Sci.*, 1994, **34**, 610.
139. P. Mata, *J. Chem. Inf. Comput. Sci.*, 1999, **39**, 1117.
140. J. Brecher, *J. Chem. Inf. Comput. Sci.*, 1999, **39**, 1226.
141. M.L. Contreras, G.M. Trevisiol, J. Alvarez, G. Arias and R. Rozas, *J. Chem. Inf. Comput. Sci.*, 1999, **39**, 1228.
142. P. Mata, *J. Chem. Inf. Comput. Sci.*, 2000, **40**, 1072.
143. C. Le Bret, *J. Chem. Inf. Comput. Sci.*, 1996, **36**, 678.
144. I. Lukovits, *J. Chem. Inf. Comput. Sci.*, 1999, **39**, 563.

145. I. Lukovits, *J. Chem. Inf. Comput. Sci.*, 2000, **40**, 361.
146. H.L. Morgan, *J. Chem. Doc.*, 1965, **5**, 107.
147. I. Lukovits, *J. Chem. Inf. Comput. Sci.*, 2000, **40**, 1147.
148. M. Deza, P.W. Fowler and V. Grishukhin, *J. Chem. Inf. Comput. Sci.*, 2001, **41**, 300.
149. G. Caporossi, P. Hansen and M. Zheng, Enumerations of Fusenes to $h = 20$, in *Discrete Mathematical Chemistry*, DIMACS Series in Discrete Mathematical and Theoretical Computer Science, Vol. 51, eds. P. Hansen, P.W. Fowler and M. Zheng, American Mathematical Society, Providence, RI, 2000, pp. 63–78.
150. G. Brinkmann, A.A. Dobrynin and A. Krause, *Comm. Math. Comput. Chem.*, 2000, **41**, 137.
151. P.W. Fowler, T. Heine and A. Troisi, *Chem. Phys. Lett.*, 1999, **312**, 77.
152. H.W. Kroto and D.R.M. Walton, *Chem. Phys. Lett.*, 1993, **214**, 353.
153. P.W. Fowler, K.M. Rogers, K.R. Somers and A. Troisi, *J. Chem. Soc. Perkin Trans. 2*, 1999, 2023.
154. B. de la Vaissiere, P.W. Fowler and M. Deza, *J. Chem. Inf. Comput. Sci.*, 2001, **41**, 376.
155. P. Hansen and N. Mladenović, *YUGOR, Yugoslav J. Operations Res.*, 1992, **2**, 3.
156. G. Brinkmann and M. Deza, *J. Chem. Inf. Comput. Sci.*, 2000, **40**, 530.
157. O.D. Friedrichs and M. Deza, More Icosahedral Symmetry Fullerenes, in *Discrete Mathematical Chemistry*, DIMACS Series in Discrete Mathematical and Theoretical Computer Science, Vol. 51, eds. P. Hansen, P. W. Fowler and M. Zheng, American Mathematical Society, Providence, RI, 2000, pp. 97–116.
158. C.M. Quinn, D.B. Redmond and P.W. Fowler, Group and Graph Theoretical Perspectives on the Structures of Large Icosahedral Cages, in *Discrete Mathematical Chemistry*, DIMACS Series in Discrete Mathematical and Theoretical Computer Science, Vol. 51, eds. P. Hansen, P.W. Fowler and M. Zheng, American Mathematical Society, Providence, RI, 2000, pp. 293–303.
159. M.S. Dresselhaus, G. Dresselhaus and P.C. Eklund, *Science of Fullerenes and Carbon Nanotubes*, Academic Press, New York, 1996.
160. H.-Y. Zhu, D.J. Klein, T.G. Schmalz, A. Rubio and N.H. March, *J. Phys. Chem. Solids*, 1998, **59**, 417.
161. P.W. Fowler and K.M. Rogers, *J. Chem. Inf. Comput. Sci.*, 2001, **41**, 108.
162. P.W. Fowler, T. Pisanski, A. Graovac and J. Žerovnik, A Generalized Ring Spiral Algorithm for Coding Fullerenes and other Cubic Polyhedra, in *Discrete Mathematical Chemistry*, DIMACS Series in Discrete Mathematical and Theoretical Computer Science, Vol. 51, eds. P. Hansen, P.W. Fowler and M. Zheng, American Mathematical Society, Providence, RI, 2000, pp. 175–188.
163. D.E. Manolopoulos, J.C. May and S.E. Down, *Chem. Phys. Lett.*, 1991, **181**, 105.
164. D.E. Manolopoulos and P.W. Fowler, *Chem. Phys. Lett.*, 1993, **204**, 1.
165. P.W. Fowler, D.E. Manolopoulos, D.B. Redmond and R.P. Ryan, *Chem. Phys. Lett.*, 1993, **202**, 371.
166. H. Terrones and M. Terrones, Geometry and Energetics of High Genus Fullerenes and Nanotubes, in *Discrete Mathematical Chemistry*, DIMACS Series in Discrete Mathematical and Theoretical Computer Science, Vol. 51, eds. P. Hansen, P.W. Fowler and M. Zheng, American Mathematical Society, Providence, RI, 2000, pp. 333–342.
167. R.B. King, Carbon Networks on Cubic Infinite Periodic Minimal Surfaces, in *Discrete Mathematical Chemistry*, DIMACS Series in Discrete Mathematical and Theoretical Computer Science, Vol. 51, eds. P. Hansen, P.W. Fowler and M. Zheng, American Mathematical Society, Providence, RI, 2000, pp. 235–248.
168. C.R. Si, *Comm. Math. Comput. Chem.*, 2000, **41**, 45.

169. D.J. Klein and N. Trinajstić, *J. Mol. Struct. (Theochem)*, 1990, **206**, 135.
170. J.R. Dias, *Phys. Chem. Chem. Phys.*, 1999, **1**, 5081.
171. J.R. Dias, *J. Chem. Inf. Comput. Sci.*, 2000, **40**, 810.
172. J.R. Dias, *J. Chem. Inf. Comput. Sci.*, 2001, **41**, 129.
173. J.R. Dias, *J. Chem. Inf. Comput. Sci.*, 2001, **41**, 686.
174. J.R. Dias, *Polycycl. Arom. Comps.*, 2000, **17**, 93.
175. D.J. Klein and T.G. Schmalz, *Phys. Rev. B*, 1990, **41**, 2244; D.J. Klein, G.E. Hite, W.A. Seitz and T.G. Schmalz, *Theor. Chim. Acta*, 1986, **69**, 409.
176. A. Graovac, O.E. Polansky and N.N. Tyutyulkov, *Croat. Chem. Acta*, 1983, **56**, 325.
177. D. Babić, A. Graovac, B. Mohar and T. Pisanski, *Discr. Appl. Math.*, 1986, **15**, 11.
178. G.G. Cash and W.C. Herndon, *Comm. Math. Comput. Chem.*, 1999, **40**, 273.
179. L.Z. Zhang and F.J. Zhang, *J. Math. Chem.*, 2000, **27**, 319.
180. I. Gutman, *J. Math. Chem.*, 1993, **12**, 197.
181. D. Babić, A. Graovac and I. Gutman, *Polycycl. Arom. Comps.*, 1995, **4**, 199.
182. S.J. Cyvin, I. Gutman and S.J. Cyvin, *Ach-Models Chem.*, 1994, **131**, 777.
183. I. Gutman and S.J. Cyvin, *J. Mol. Struct. (Theochem)*, 1993, **107**, 85.
184. I. Gutman, Y.N. Yeh, S.L. Lee, H. Hosoya and S.J. Cyvin, *Z. Naturforsch.*, 1994, **49a**, 1053.
185. A. Graovac, M. Juvan, B. Mohar and J. Žerovnik, *Croat. Chem. Acta*, 1999, **72**, 853.
186. J.R. Dias, *Int. J. Quantum Chem.*, 1999, **74**, 721.
187. O.E. Polansky and N.N. Tyutyulkov, *Comm. Math. Comput. Chem.*, 1977, **3**, 149.
188. A. Graovac and D. Babić, *Z. Naturforsch.*, 1985, **40a**, 66.
189. C.D. Lin and G.Q. Fan, *J. Chem. Inf. Comput. Sci.*, 1999, **39**, 782.
190. D.J. Klein and N. Trinajstić, *Pure Appl. Chem.*, 1999, **39**, 782.
191. C.D. Lin, *J. Chem. Inf. Comput. Sci.*, 2000, **40**, 778.
192. I. Gutman, C. Rucker and G. Rucker, *J. Chem. Inf. Comput. Sci.*, 2001, **41**, 739.
193. S. Marković, Z. Marković and R.I. McCrindle, *J. Chem. Inf. Comput. Sci.*, 2001, **41**, 112.
194. E. Estrada, *J. Chem. Inf. Comput. Sci.*, 1996, **36**, 844.
195. I. Gutman, S. Klavžar, M. Petkovšek and P. Zigert, *Comm. Math. Comput. Chem.*, 2001, **43**, 49.
196. G. Cash, *J. Chem. Inf. Comput. Sci.*, 2000, **40**, 1203, 1207.
197. K. Nitta, *J. Math. Chem.*, 1999, **25**, 133.
198. M.D. Tissandier, S.J. Singer and J.V. Coe, *J. Phys. Chem. A*, 2000, **104**, 752.
199. F. Harary and E. Palmer, *Bull. Acad. Polon. Sci. Ser. Sci. Math. Astronom. Phys.*, 1966, **14**, 25.
200. F. Harary and E. Palmer, *Can. J. Math.*, 1966, **18**, 853.
201. S. McDonald, L. Ojamäe, and S.J. Singer, *J. Phys. Chem.*, 1998, **102**, 2824.
202. T.P. Radhakrishnan and W.C. Herndon, *J. Phys. Chem.*, 1991, **95**, 10609.
203. P. Vismara and C. Laurencio, An Abstract Representation for Molecular Graphs, in *Discrete Mathematical Chemistry*, DIMACS Series in Discrete Mathematical and Theoretical Computer Science, Vol. 51, eds. P. Hansen, P.W. Fowler and M. Zheng, American Mathematical Society, Providence, RI, 2000, pp. 343–366.
204. A. Dietz, C. Fiorio, M. Habib and C. Laurencio, Representation of Stereochemistry using Combinatorial Maps, in *Discrete Mathematical Chemistry*, DIMACS Series in Discrete Mathematical and Theoretical Computer Science, Vol. 51, eds. P. Hansen, P.W. Fowler and M. Zheng, American Mathematical Society, Providence, RI, 2000, pp. 117–128.
205. Y. Xu and M. Johnson, *J. Chem. Inf. Comput. Sci.*, 2001, **41**, 181.

206. *Topological Indices and Related Descriptors in QSAR and QSPR*, ed. J.E. Devillers and A.T. Balaban, Gordon and Breach Science Publishers, NY, 1999.
207. R. Todeschini and V. Consonni, *Handbook of Molecular Descriptors*, Wiley-VCH, Mannheim, 2000.
208. *QSAR/QSPR Studies by Molecular Descriptors*, ed. M.V. Diudea, Nova Science Publishers, NY, 2000.