Chemical Graphs. 43. FORTRAN IV Program for Computing the Numbers of General Cubic Graphs on p Vertices†

ALEXANDRU T. BALABAN*

Department of Organic Chemistry, The Polytechnic, Bucharest, Roumania

ROMUL VANCEA

Regional Computing Centre, 5800 Suceava, Roumania

IOAN MOTOC

Centre of Chemistry, 1900 Timisoara, Roumania

STEFAN HOLBAN

Computer Department, The Polytechnic, Timisoara, Roumania

Received January 14, 1985

A computer program is described for calculating the numbers $N(p)$ of general cubic graphs on p vertices. Numerical results are given for $p = 2, 4, 6, 8, 10$, and 12. A recursive algorithm was devised for obtaining the numbers $D(p)$ of disconnected cubic general graphs, making use of the Vandermonde convolution formula. By difference, the numbers $K(p) = N(p) - D(p)$ of connected general cubic graphs are obtained; these numbers $K(p)$ coincide with data (obtained previously for chemical purposes) of connected general cubic graphs with p = 2–10, which had been obtained by a different, constructive method. Numbers $P(p)$ of connected planar cubic multigraphs with p = 4–12 are indicated; these graphs represent constitutional formulas for valence isomers of [p]annulenes.

INTRODUCTION

The enumeration of cubic graphs has remained on the list of unsolved graph-theoretical problems for a long time,[1] but now the problem appears to be solved,[2] because a formula (see eq 1) was demonstrated. However, no numerical data have yet been obtained from this formula. In connection with the chemical problem of the valence isomers of annulenes, $(CH)_p$, whose constitutional (structural) formulas are connected cubic (trivalent) multigraphs where each vertex stands for a CH group, we were interested in the enumeration of cubic multigraphs and general graphs. All valence isomers of annulenes that have chemical significance correspond to connected planar cubic multigraphs. A constructive algorithm was found,[3] and it was applied for finding the constitutional formulas (hence the numbers) of connected cubic multigraphs and general graphs with p = 2, 4, 6, 8, and 10 vertices.[3–5] (According to a well-known theorem, the order p of cubic graphs is an even integer because the sum of the vertex degrees equals twice the number q of edges, $3p = 2q$. For chemical compounds with odd numbers of carbon atoms, general cubic graphs with one loop were used;[5] the loop and its attached vertex symbolize a heteroatom with vertex degree equal to 2.) In a subsequent paper, the constitutional formulas (and hence the numbers) of connected planar cubic multigraphs with p = 12 vertices were displayed.[6] The numbers of simple connected cubic graphs with 12 or more vertices were subsequently discussed by Bussemaker et al.,[7] Faradzhev,[8] Sasaki,[9] Robinson,[9] and Khvorostov.[9] Thus, although earlier work[3–6] gave constructive and numerical data, it did not produce an algorithm suitable for the desired numerical solution. That, however, is taken care of in the present report by the computer implementation of formula 1.

This present paper describes a computer program for finding the numbers $N(p)$ of general cubic graphs with p vertices (connected and disconnected, multigraphs and loop-graphs, planar and nonplanar). The program is based on the formula[2]

$$N(p) = Z(S_p[S_3]) \cap Z(S_{3p/2}[S_2]) \tag{1}$$

where Z is the cycle index and S_p denotes the symmetric permutation group of the $p!$ permutations of p objects.

It is known that[10]

$$Z(S_n) = \frac{1}{n!}\sum_t \prod_i \frac{n!x_i^{j_i}}{\prod_k k^{j_k} j_k!} \tag{2}$$

where $\sum_t$ has $p(n)$ terms, i.e., all partitions of n, $i = 1$–n, and $k = 1$–r. For a given partition, $x_1^{j_1}x_2^{j_r}...x_r^{j_r}$, where

$$1j_1 + 2j_2 + ... + rj_r = n \tag{3}$$

the denominator of (2) is therefore

$$\prod_{k=1}^{r} k^{j_k} j_k! \tag{4}$$

It is also known that one may replace $x_i^{j_i}$ by $[Z(S_m)]_{(i)}^{j_i}$ in formula 2:

$$Z(S_n[S_m]) = Z(S_n, Z(S_m)) = \frac{1}{n!}\sum_t \prod_i \frac{n![Z(S_{mi})]^{j_i}}{k^{j_k} j_k!} \tag{5}$$

For applying formula 1, one has to compute separately $Z(S_p[S_3])$ and $Z(S_{3p/2}[S_2])$ and to add the direct products of all elements common to both of them. Thus, let an element of $Z(S_p[S_3])$ be $Ax_1^{j_1}x_2^{j_2}...x_r^{j_r}$ and an element of $Z(S_{3p/2}[S_2])$ be $Bx_1^{j_1}x_2^{j_2}...x_r^{j_r}$. Then, N will contain the element

$$ABx_1^{j_1}x_2^{j_2}...x_r^{j_r} \cap x_1^{j_1}x_2^{j_2}...x_r^{j_r} = AB\prod_{k=1}^{r} k^{j_k} j_k! \tag{6}$$

As an example for $p = 2$, $N(2)$ is the intersection:

$$N(2) = Z(S_2[S_3]) \cap Z(S_3[S_2]) \tag{7}$$

$$Z(S_2) = (x_1^2 + x_2)/2 \tag{8a}$$

† The preceding paper in this series is Balaban, A. T.; Biermann, D.; Schmidt, W. *Nouv. J. Chim.* 1985, 9, 443–449.

0095-2338/86/1626-0072$01.50/0

$$Z(S_3) = (x_1^3 + 3x_1x_2 + 2x_3)/6 \tag{8b}$$

therefore

$$Z(S_2[S_3]) = \frac{[Z(S_3)]_{(1)}^2 + [Z(S_3)]_{(2)}}{2} = \frac{(x_1^3 + 3x_1x_2 + 2x_3)^2/6^2 + (x_2^3 + 3x_2x_4 + 2x_6)/6}{2} = (x_1^6 + 9x_1^2x_2^2 + 4x_3^2 + 6x_1^4x_2 + 18x_2x_4 + 12x_6 + 6x_2^3 + 4x_1^3x_3 + 12x_1x_2x_3)/72 \tag{9}$$

and

$$Z(S_3[S_2]) = \{[Z(S_2)]_{(1)}^3 + 3[Z(S_2)]_{(1)}[Z(S_2)]_{(2)} + 2[Z(S_2)]_{(3)}\}/6 = [(x_1^2 + x_2)^3/8 + 3(x_1^2 + x_2)(x_2^2 + x_4)/4 + 2(x_3^2 + x_6)/2]/6 = (x_1^6 + 9x_1^2x_2^2 + 8x_3^2 + 3x_1^4x_2 + 7x_2^3 + 6x_2x_4 + 8x_6 + 6x_1^2x_4)/48 \tag{10}$$

It may be seen that all but the last two elements in (9) and all but the last element in (10) are common both to (9) and to (10); therefore, by substituting in (7) according to (6), we obtain

$$N(2) = (x_1^6 \cap x_1^6 + 81x_1^2x_2^2 \cap x_1^2x_2^2 + 32x_3^2 \cap x_3^2 + 18x_1^4x_2 \cap x_1^4x_2 + 42x_2^3 \cap x_2^3 + 108x_2x_4 \cap x_2x_4 + 96x_6 \cap x_6)/72{\cdot}48 = (6! + 81{\cdot}2!{\cdot}2^2{\cdot}2! + 32{\cdot}3^2{\cdot}2! + 18{\cdot}4!{\cdot}2 + 42{\cdot}2^3{\cdot}3! + 108{\cdot}2{\cdot}4 + 96{\cdot}6)/72{\cdot}48 = 2 \tag{11}$$

There are, as seen above, seven common terms in $N(2)$; however, the number of these common terms increases very fast with p: 28 for $N(4)$, 87 for $N(6)$, 321 for $N(8)$, etc. Therefore, in order to implement a numerical use of relation (1), a computer program is necessary.

DESCRIPTION OF THE PROGRAM

Name: Isoval.
Computer: Felix C-1024.
Language: FORTRAN IV.
Internal subroutine: none.
Memory: 900 kilobytes for $p = 12$.
Cards: 441.

The structure of the program is presented in Figure 1. Segment RO calls the MELIFORA subroutine. The segment R of the program consists of subroutines RAD, COD, ADUN, MUT, PUTERE, DECAL, INMUL, and FACTNUM or FACTNUM1. Then, the program branches into three segments. Segment B contains subroutine PART; segment C contains ADUN1 and INTER; segment D contains subroutines READ, MELIFORA, and MELIFORD. The last two subroutines are written in Assembler language for the Felix computer and on other computers should be replaced by stating the matrix dimensions. Table I presents in more detail both the subroutines composing each segment and the mode by which a subroutine calls other subroutines (subroutine interlinking). The block diagram (Figure 2) represents the logical scheme of the Isoval program.

PROGRAM INPUT AND OUTPUT DATA

The first card contains in format 2I3 the numbers NP3 and NP4 according to formula 12, which translates formula 1:

$$\text{(NP1 or NP2) and (NP3 or NP4)} \tag{12}$$

Formula 12 includes notations *or* and *and* because although we operate with matrices and perform additions, multiplications, and intersections, the operations symbolized by *and* or by *or* do not have a direct correspondence in matrix calculus or Boolean algebra.

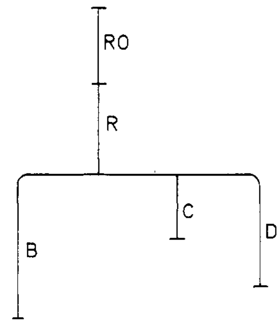


Figure 1. Structure of the Isoval program.

Table I. Structure of the Isoval Program

segment	subroutine	called subroutine(s)
RO	MELIFORA	none
	RAD	READ, PART, INMUL
R	RAD	READ, PART, INMUL
	COD	MUT, PUTERE, INMUL, ADUN, ADUN1, INTER
	ADUN	COD
	MUT	none
	PUTERE	INMUL
	DECAL	none
	INMUL	none
	FACTNUM	none
	FACTNUM1	none
B	PART	FACTNUM
C	ADUN1	none
	INTER	FACTNUM1
D	READ	none
	MELIFORA	none
	MELIFORD	none

Print-out tables give the four partitions of NP1, NP2, NP3, and NP4, followed by the final result presented as $V = ...$ (no. of graphs). The listing of the Isoval program is available as supplementary material (see paragraph at end of paper regarding supplementary material).

STRUCTURE OF THE PROGRAM

Mode of Representation for Numerical Data. Since the program operates with sparse matrices of large dimensions, in order to save memory space we had recourse to equivalent matrices whose entries b_{ik} cumulate both the value and the position of the nonzero entries a_{ij} in the sparse matrix:

$$b_{ik} = 100j + a_{ij}$$

Figure 3 illustrates this conversion of a sparse matrix $\mathbf{A} = \|a_{ij}\|$ into matrix $\mathbf{B} = \|b_{ik}\|$. This conversion $\mathbf{A} \rightarrow \mathbf{B}$ results in a reduction by approximately 60% of the active memory area that is utilized, without increasing significantly the computing time necessary for coding and decoding the entries: the number s of the columns in $\mathbf{A}$ is reduced to t in $\mathbf{B}$, equalling the maximum number of nonzero entries in any row of $\mathbf{A}$.

Organization of Data in the Program. The following matrices were employed: six working matrices $\mathbf{M}_i$ with integer entries and dimension 500 × 10; six working matrices $\mathbf{V}_i$ with double-precision entries and dimension 500 × 10; one matrix for storing integer results **MP**, with dimension 3600 × 10; two matrices for storing double-precision results **VP** and **VS**, with dimension 3600 × 1.

In order to speed up calculations, all these matrices are stored totally and simultaneously in the memory. This fact imposed the used of specially developed routines (MELIFORA and MELIFORD) because the Felix computer does not allow program segments with requirements larger than 64 kilobytes, taking into account that a single $\mathbf{M}_i$ matrix exceeds 90 kilo-

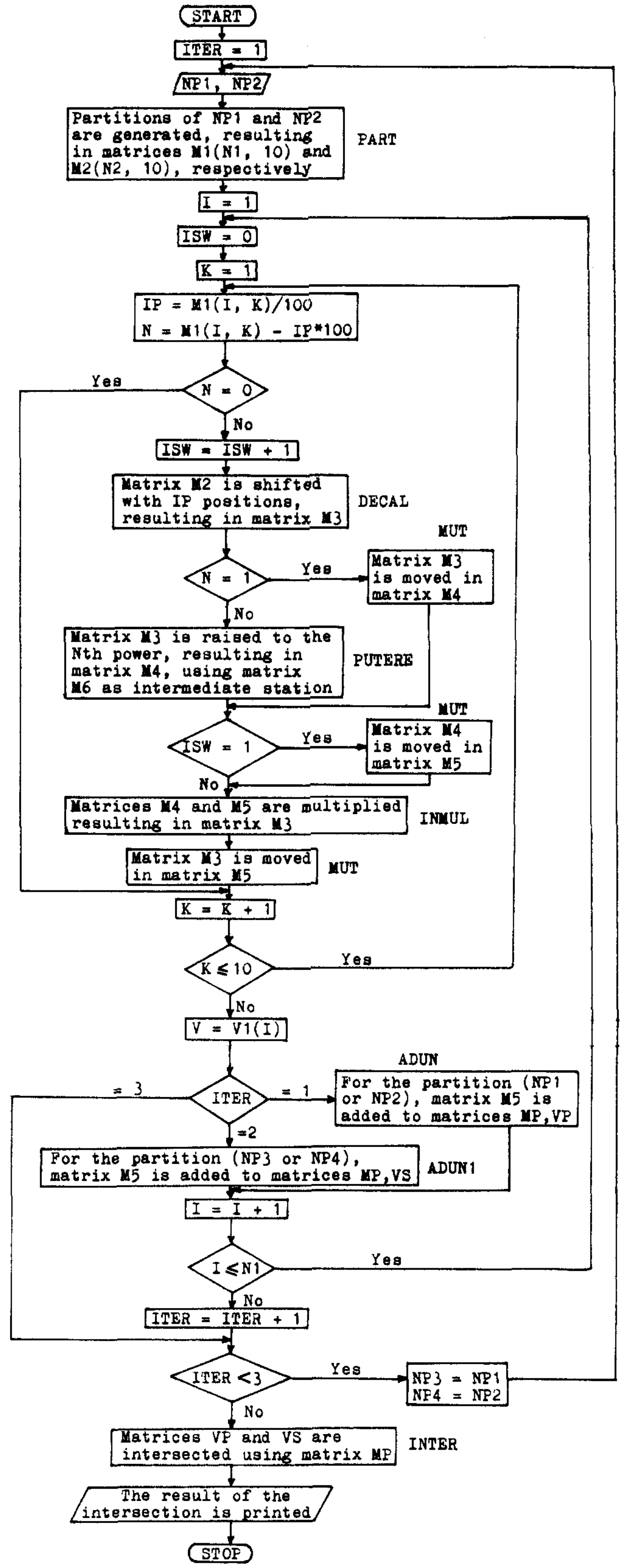


Figure 2. Block diagram of the Isoval program.

bytes. On other computers such as CDC or IBM, these routines are not necessary; the structure of the program remains the same.

RESULTS

Table II presents the results known earlier, together with the results newly obtained with the help of Isoval. In order to compare the new results with the previous ones, it is necessary to subtract the disconnected general cubic graphs, because in chemistry only the connected cubic graphs are significant. The total execution time was 3 h and 26 min for the last value $N(12)$ from Table II.

ENUMERATION OF DISCONNECTED CUBIC GRAPHS

There exist no disconnected general cubic graphs with two

Table II. Numbers of Cubic Graphs

	general graphs			connected multigraphs			connected simple graphs			
no. of vertices, p	disconnected, $D(p)^a$	connected, $K(p)^b$	total, $N(p)^a$	planar, $P(p)^c$	nonplanar, $U(p)^d$	total, $M(p)^c$	planar, $A(p)^e$	nonplanar, $R(p)^e$	total, $S(p)^e$	loop-graphs, $L(p)^f$
2	0	2	2	1	0	1	0	0	0	1
4	3	5	8	2	0	2	1	0	1	3
6	14	17	31	5	1	6	1	1	2	11
8	69	71	140	17	3	20	3	2	5	51
10	334	388	722	71	20	91	9	10	19	297
12	1847	2592	4439	357	152	509	32	53	85	2086

[a] Present results. [b] See formula 13; for $p = 10$, see references 4 and 5. [c] See references 3 and 4 for $p = 10$ and reference 6 for $p = 12$. [d] See formula 22. [e] See references 7–10. [f] See formula 21.

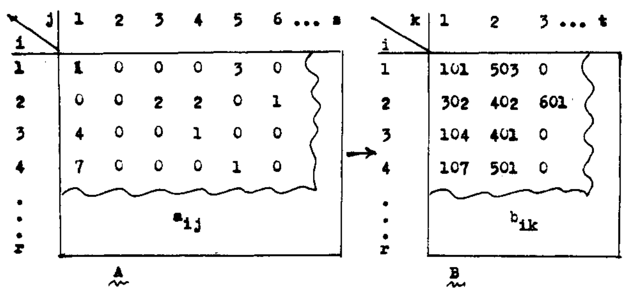
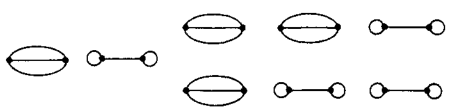


Figure 3. Conversion of a sparse matrix **A** into matrix **B**.

vertices. Since there exist exactly two connected general cubic graphs with two vertices, I and II, the three nonisomorphic



I II III IV V

combinations III–V constitute the disconnected general cubic graphs on four vertices. The disconnected general cubic graphs on six vertices are either combinations of a copy each from the set {I, II} and the set of five connected general cubic graphs on four vertices or combinations of three copies from the set {I, II}; the former group of combinations affords $2 \times 5 = 10$ disconnected graphs, and the latter group of combinations affords four disconnected graphs yielding a total of 14 disconnected cubic graphs on six vertices.

We proceed to devise a recurrence formula able to split the number $N(p)$ of general cubic graphs on p vertices into connected, $K(p)$, and disconnected, $D(p)$ graphs according to

$$N(p) = D(p) + K(p) \tag{13}$$

We shall reformulate part of our task as follows:

Problem. Let us have sets of f_p nonidentical objects (e.g., connected general graphs on p vertices, i.e., $f_2 = 2, f_4 = 5$, etc.). In how many distinct ways X can we combine k unordered objects (the objects need not be distinct)? Another example is of an alphabet of f letters; if a k-letter word is defined as an unordered sequence of k letters, multiple letters being allowed, then we may ask how many distinct k-letter words X are possible. (If a k-letter word would be defined as an ordered sequence of k letters, the solution would be, evidently, $X = f^k$.)

Solution

$$X = C^k_{f+k-1} = \binom{f+k-1}{k} = \frac{(f+k-1)!}{k!(f-1)!} \tag{14}$$

Proof. We have

$$X = C^{k-1}_{k-1}C^1_f + C^{k-2}_{k-1}C^2_f + C^{k-3}_{k-1}C^3_f + \ldots \tag{15}$$

Since the first product gives the number of words consisting of one kind of letters, the second product gives the number of words consisting of two kinds of letters, and so on. According to Vandermonde's convolution formula 16, irrespective

$$C^k_{f+k-1} = \sum_{i=0}^{\min(f,k)} C^{k-i-1}_{k-1}C^{i+1}_f \tag{16}$$

of whether k is greater than or less than f, relations 14 and 15 are shown to be equal. This completes the proof.

Coming back to the number $D(p)$ of disconnected general cubic graphs, let the number of partitions of $p/2$ be Y. Let one of these partitions be

$$x_1^{j_1}x_2^{j_2}...x_r^{j_r} \tag{17}$$

where $1j_1 + 2j_2 + ... + rj_r = p/2$; cf. also (3). We multiply by 2 and obtain

$$2j_1 + 4j_2 + ... + 2rj_r = p \tag{18}$$

We symbolize the disconnected graphs corresponding to this partition of p by

$$K_2^{j_1}K_4^{j_2}...K_{2r}^{j_r} \tag{19}$$

We apply now to each of the $Y - 1$ partitions consisting of at least two terms relation 14 in the form

$$X = \prod_{i=1}^{r} C^{j_i}_{K(2i)+j_i-1} \tag{20}$$

We may add (19) after the number (20) in order to indicate the type of disconnected graphs. This recurrence algorithm for disconnected general cubic graphs affords for $p = 2, 4, ..., 12$ the following results:

(i) For $p = 2$, $N(2) = 2$, $D(2) = 0$, and $K(2) = 2$.

(ii) For $p = 4$, we wish to obtain $D(4)$. There is only $Y - 1$ = one partition of 4/2 with at least two terms, namely, $1 + 1$, symbolized by 1^2, according to (3) or (17). Multiplying by 2, we obtain the partition of 4 as $2 + 2$, i.e., two graphs on two vertices symbolized by K_2^2. According to (14), where $k = f = 2$, or to (19) and (20), we obtain $X = 3K_2^2$; hence, $D(4) = 3$ (graphs III–V). Since the program gives $N(4) = 8$, we obtain by difference $K(4) = 5$.

(iii) For $p = 6$, we wish to obtain $D(6)$. There are $Y - 1$ = two partitions of 6/2 with at least two terms, namely, 1^3 and 1, 2; therefore, we have disconnected graphs of the form K_2^3 and K_2K_4. In the former case, according to (19) and (20), we obtain $X = 4K_2^3$, and in the latter case, we obtain $X = 10K_2K_4$; hence, $D(6) = 4 + 10 = 14$ (graphs III–V). By difference, it results that $K(6) = 17$.

(iv) For $p = 8$, analogously we obtain $5K_2^4$, $15K_4^2$, $15K_2^2K_4$, and $34K_2K_6$, totaling $D(8) = 69$.

(v) For $p = 10$, we obtain $6K_2^5$, $20K_2^3K_4$, $30K_2K_4^2$, $51K_2^2K_6$, $142K_2K_8$, and $85K_4K_6$, totaling $D(10) = 334$.

(vi) For $p = 12$, we obtain $7K_2^6$, $25K_2^4K_4$, $68K_2K_6^3$, $45K_2^2K_4^2$, $45K_4^3$, $170K_2K_4K_6$, $213K_2^2K_8$, $153K_6^2$, $355K_4K_6$, and $776K_2K_{10}$, totaling $D(12) = 1847$. The next value is $D(14) = 10\,995$, etc.

DISCUSSION OF THE RESULTS

Previous papers[3–6] gave, in addition to constitutional formulas, the numbers of connected cubic multigraphs with p =

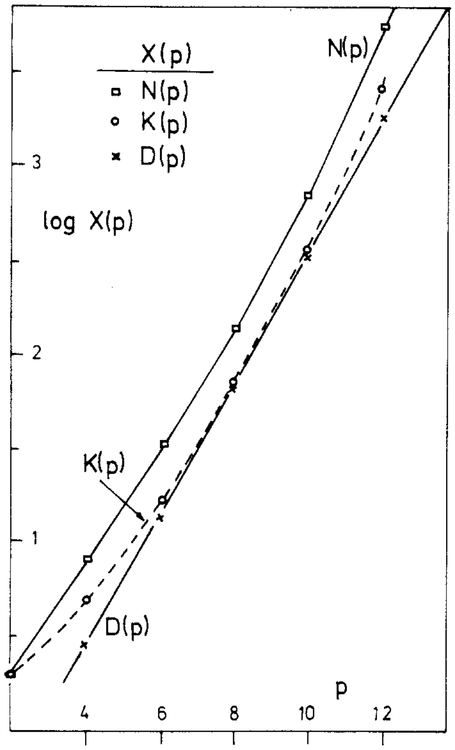


Figure 4. Semilogarithmic plot of $N(p)$, upper full line, $K(p)$, broken line, and $D(p)$, lower line, vs. p.

2–10, of connected cubic planar multigraphs with $p = 12$ and of connected general cubic graphs with $p = 2$–10. The numbers thus determined previously coincide exactly with those determined with the help of the Isoval program, and presented in Table II.

The numbers of general cubic graphs on 12 vertices, $N(12)$, $D(12)$, and $K(12)$, are now determined for the first time. From previously determined numbers of connected multigraphs $M(p)$ and connected planar multigraphs $P(p)$,[4–6] one can determine by difference the numbers of connected loop-graphs $L(p)$ and of connected nonplanar multigraphs $U(p)$:

$$L(p) = K(p) - M(p) \quad (21)$$

$$U(p) = M(p) - P(p) \quad (22)$$

Numbers of simple cubic graphs [total number $S(p)$, planar ones $A(p)$ and nonplanar ones $B(p)$] are taken from literature data.[7–9] A chemical use of loop-graphs was indicated here earlier in the text.[5]

As indicated in Figure 4, the numbers $N(p)$ of general cubic graphs increase a little faster than exponentially vs. p. The numbers $K(p)$ and $D(p)$ of connected and disconnected general cubic graphs, respectively, increase in a parallel analogous fashion, with the former being higher than the latter: log $D(p)$ increases almost linearly vs p in the range $4 \leq p \leq 12$. We obtain $\log D(p) \approx 0.34782p - 0.93206$ by the method of least squares for this range of p values; the correlation coefficient is $r^2 = 0.99966$, yet despite this excellent value, the trend is the same as for $N(p)$ and $K(p)$. Therefore, extrapolations of Figure 4 provide lower bounds for $N(p)$, $K(p)$, and $D(p)$.

At present, we are expanding the program and the matrix dimensions for accomodating the calculation of $N(p)$ for $p > 12$. Recent reviews[11–13] have highlighted the spectacular progress of synthetic work in the area of valence isomers of [8]- and [10]annulene[11,12] as well as of [12]annulene.[12,13]

The present data validate previously published numbers of connected cubic multigraphs with up to 10 vertices.[3–5] There is complete agreement between the structures and number (226) of nonseparable (bridge-free, isthmus-free) connected planar multigraphs with 12 vertices; these data were obtained by two independent approaches.[6,18] The chemically relevant number is $P(12) = 357$ planar connected multigraphs (separable and nonseparable).

ACKNOWLEDGMENT

Thanks are addressed to Dr. D. H. Smith for having provided the print-out coding of nonseparable (planar and nonplanar) cubic multigraphs with 12 vertices.

Supplementary Material Available: Listing of Isoval program (14 pages). Ordering information is given on any current masthead page.

REFERENCES AND NOTES

(1) Harary, F. *Publ. Math. Inst. Hung. Acad. Sci.* **1960**, *5A*, 63–95.
(2) Harary, F.; Palmer, E. M. *Graphical Enumeration*; Academic: New York, 1973; p 174. Read, R. C. *J. London Math. Soc.* **1959**, *34*, 417–436. Ibid. **1960**, *35*, 344–351.
(3) Balaban, A. T. *Rev. Roum. Chim.* **1966**, *11*, 1097–1116. Erratum, Ibid. **1967**, *12*, no. 1, last page.
(4) Balaban, A. T. *Rev. Roum. Chim.* **1970**, *15*, 463–485.
(5) Balaban, A. T. *Rev. Roum. Chim.* **1974**, *19*, 1611–1619. Erratum, Ibid. **1978**, *23*, 311.
(6) Balaban, A. T. *Rev. Roum. Chim.* **1972**, *17*, 865–881.
(7) Bussemaker, F. C.; Cobeljic, S.; Cvetkovic, D. M.; Seidel, J. J. *J. Comb. Theory*, **1977**, *B23*, 234–235. "Computer Investigation of Cubic Graphs"; T. H. Report 76, 1976; WSK-01, Technology University: Eindhoven, The Netherlands.
(8) Faradzhev, N. A. *Usp. Mat. Nauk* **1976**, *31*, 246–248.
(9) Abe, H.; Okuyama, T.; Fujiwara, I.; Sasaki, S. *J. Chem. Inf. Comput. Sci.* **1984**, *24*, 220–229. Khvorostov, P. V. "Mashinnye Metody Obnaruzhenya Zakonomernostei, Analiza Struktur i Proektirovanya". *Vychisl. Sist.* **1982**, *92*, 80–141. Robinson, R. W.; Wormald, N. C. *J. Graph Theory* **1983**, *7*, 463–467.
(10) Kokhov, V. A.; Kuznetsov, I. O.; Lazarev, V. A. "Teoria Grafov". *Tr. Mosk. Energ. Inst.* **1975**, *No. 250*, 113–122.
(11) Randic, M. *Acta Crystallogr., Sect. A*: *Cryst. Phys., Diffr., Theor. Gen. Crystallogr.* **1978**, *A34*, 275–282.
(12) Balaban, A. T.; Davies, R. O.; Harary, F.; Hill, A.; Westwick, R. *Aust. J. Math.* **1970**, *11*, 207–215.
(13) Harary, F. In *A Seminar on Graph Theory*; Harary, F., Ed.; Holt, Rinehart and Winston: New York, 1967; p 26. *Graph Theory in Theoretical Physics*; Academic: New York, 1967; p 8.
(14) Balaban, A. T.; Banciu, M. *J. Chem. Educ.* **1984**, *61*, 766–770.
(15) Balaban, A. T.; Banciu, M.; Ciorba, V. *Annulenes, Benzo-, Hetero-, Homo-Derivatives and Their Valence Isomers*; CRC: Boca Raton, FL, 1986.
(16) Banciu, M.; Popa, C.; Balaban, A. T. *Chem. Scr.* **1984**, *24*, 28–37.
(17) Lederberg, J. "DENDRAL-64. A System for Computer Construction, Enumeration and Notation of Organic Molecules as Tree Structures and Cyclic Graphs. Part II. Topology of Cyclic Graphs"; NASA Report CR-68898, 15 Dec 1965; NASA: Washington, DC.
(18) Carhart, R. E.; Smith, D. H.; Brown, H.; Sridharan, N. S. *J. Chem. Inf. Comput. Sci.* **1975**, *15*, 124–130.