Journal of Mathematical Chemistry 3(1989)193–115

EMBEDDING FREQUENCIES OF TREES

R.D. POSHUSTA and M.C. McHUGHES

Chemical Physics Program, Washington State University, Pullman, WA 99164-4630, USA

Received 16 May 1988
(in final form 3 November 1988)

Abstract

A graph γ is said to be embedded in a graph Γ if γ is isomorphic to a subgraph of Γ. The embedding frequency for γ in Γ, $N(\Gamma, \gamma)$, is the number of different subgraphs of Γ to which γ is isomorphic. We use a computer program to calculate the embedding frequencies of subtrees within trees. We compute $N(\Gamma, \gamma)$ for trees through 10 vertices and present the results in tabular form. When trees are partially ordered by valence class, their subtrees lie in corresponding order; we give a formal proof of this subtree embedding property. The structure of the embedding relation is exhibited in a topological picture of the zeta function showing the non-zero values of $N(\Gamma, \gamma)$.

1. Introduction

Graph theoretic cluster expansions of physical, chemical, and even biological activity have received attention recently [1–7]. Cluster expansions are inspired by the intuitive belief that properties of a composite molecule arise from contributions made by its constituent functional groups: "the whole is equal to the sum of its parts". One such cluster expansion method is based on the number of times a given cluster appears as a subgraph within the larger graph representing the molecule of interest [3,4]. This approach requires the matrix of embedding frequencies for all graphs within each other.

Subgraph enumeration belongs to a long tradition of chemical graph theoretic counting problems. Numbers of isomers of various empirical formulas were first to be counted [8–10]. Cayley, for example, found a systematic method for counting the number of isomeric saturated hydrocarbons [8]. Polya [9] showed how to count the numbers of isomers obtainable by substitution on a more or less symmetrical backbone (e.g. the numbers of $C_{6+n}H_{6+2n}$ isomers derived by substituting aliphatic groups on benzene). Enumeration of all trees with up to 80 vertices along with many

other graph theoretic and chemical graph theoretic enumerations may be found in the book by Knop, Müller, Szymanski and Trinajstić [11].

In contrast to the enormous literature of graphical enumeration [11,12] and chemical isomer enumeration, there is little published work on enumerating subgraphs. It has been remarked [5] that "counting substructures is generally a much harder problem than counting graphs or isomers". Algorithms have been devised to count the total number of subtrees of a given tree [13] (regardless of repetitions). In other cases, certain special subgraphs have been counted [14]. Kier and Hall [7] have previously enumerated certain subgraphs, which they then apply to properties of various compounds; unfortunately, they published only an incomplete enumeration algorithm and a partial list of embedding frequencies. Gordon and Kennedy [15], whose work most closely resembles ours, enumerated subgraphs of alkanes with up to 8-carbon backbones, which they then applied to various properties; however, their computed embedding frequencies and their algorithms were not published. Transfer matrix methods have been used to count certain subgraphs in hexagonal lattices [16].

Our present work is confined to all trees with up to ten vertices and nine edges. There are 201 such trees. We report the number of times that each tree appears as a subtree within the others. These data may prove useful for studying the mathematical properties of embedded subgraphs as well as for physical applications. The present paper is devoted to the enumeration problem and certain mathematical properties of the embedded frequencies evident in the data. In a later paper, we make application of these results to physical properties. Details of the computer programs we use will be published elsewhere.

Efficient computer representations for graphs have been studied extensively, culminating in various "codes" to specify each N-vertex *tree* by an N-tuple of integers. We find these "code" or "N-tuple" representations of trees to be very helpful for subtree enumeration. Although we opt to use Read's walk around valency code [17], the closely related N-tuple defined by Knop [18] is equally useful. Thus, the list of trees and subtrees is recorded in the form of their codes or N-tuples. This saves space and time for table look-up during enumeration. However, our initial naive hope that codes would simplify the computation of repetition frequencies was frustrated. We were able to perform the enumerations only by means of computer programs employing the adjacency matrices as intermediates. Subtree embedding frequencies are a far from trivial enumeration problem.

In the course of enumerating subtrees, we have observed certain previously unrecognized regularities. Such regularities imply new interpretations which may help our understanding of chemical structure as well as the purely mathematical structure of the graph embedding relation. In addition to the tabulated values of embedding frequencies, we find a topological map of these frequencies to be a powerful tool for searching out new structural relations and concepts; we publish it here

in the expectation that others will also find it useful. The map led us to order trees according to their valence classes, a mathematical expression of the intuitive "branching" concept. In this order, the embedding frequency map shows a more regular structure than other orderings we have tried. We have translated the most conspicuous structure of the map into a theorem relating the order of valence classes of subgraphs to the order of the graphs in which they are embedded.

2. Definitions

Let $\Gamma = (V, E)$ denote a graph whose vertex set is V and whose edge set is E. Of course, Γ can be drawn as a collection of vertex points connected by edge lines. The vertices are conveniently numbered: $V = \{i, \text{for } i = 1, 2, 3, \ldots, n\}$. The number of vertices, $n = |V|$ or $n = |\Gamma|$, is called the *order of* the graph. The edges consist of non-ordered pairs of vertices: $E = \{(i, j), \ldots,\}$.

Most of our results apply to graphs without cycles. Such graphs are called *trees*. Some of our discussion applies to graphs in general and not just to trees. Statements about "graphs" should be interpreted as applying in general, while those about trees may only apply to graphs without cycles.

A *subgraph* of the graph Γ is any graph γ whose vertex and edge sets are contained in those of Γ: $V(\gamma) \subseteq V(\Gamma)$ and $E(\gamma) \subseteq E(\Gamma)$. The null graph ϕ and the graph Γ are "improper" subgraphs of Γ. We are concerned only with *connected* graphs and connected subgraphs. A graph is said to be connected if there is at least one path between every *pair* of vertices. A *path* from i to j is a sequence of edges $(i, k), (k, l), \ldots, (m, n), (n, j)$ commencing with i and terminating with j. The *length* of such a path is the number of *edges* it contains. The length of the shortest path from i to j is the *distance* from i to j. The *diameter* of a graph is the largest distance existing in that graph.

One mathematical representation of a graph is its adjacency matrix, familiar to chemists from Hückel MO theory. One disadvantage of this representation for computers is the large space requirements: the number of elements in the adjacency matrix increases as the square of the order of the graph. The adjacency matrix has many uses. For example, $[A^n]_{ij}$ is the number of length n paths between vertices i and j. Also useful is the distance matrix: $[D]_{ij}$ is the length of the shortest path between vertices i and j.

The adjacency matrix of a graph is not unique. There are many ways to number the vertices and hence to order the rows and columns of the adjacency matrix. Two graphs which differ only in the numbering of their vertices are *isomorphic*. Two graphs Γ and Γ' are said to be isomorphic if there is a one-to-one mapping between their vertex sets V and V' which preserves the edges; that is, if (i, j) is an edge of Γ, then the image (i', j') is an edge of Γ'.

Each vertex i in Γ has a unique *valence*: the number of edges terminating at i, denoted by v_i. For applications to chemistry, the valence of any vertex is usually

limited, e.g. 4-valent carbon atoms; for our work, we make no restriction on valence. A *leaf* is any one-valent vertex: such vertices form the *exterior* of a tree. The process of removing a leaf and its edge is called *pruning*.

3. Valence classes

We find it useful to classify graphs by an equivalence relation which associates them in classes we call *valence classes*. The same equivalence has also been used by Randić [19] and Ruch [20] to characterize or quantify graphical "branching". The collection of valences for all vertices of Γ is the valence class of Γ and is denoted $v(\Gamma) = \{v_i, \text{ all } i \in \Gamma\}$. As the name suggests, the valence class quantifies the amount and type of branching in the graph (although not the relative order of branches). The order of terms in the valence class is unimportant. Common practice is to arrange the valences in non-descending order, e.g. (1,1,1,1,2,2,4) is the valence class for the graphs

We write the valence class by exhibiting the valences:

$$\alpha = \{1,1,\ldots,1\,(\alpha_1 \text{ times}),\ 2,2,\ldots,2(\alpha_2 \text{ times}), \text{etc.}\},$$

or, more briefly,

$$\alpha = \{1^{\alpha_1}, 2^{\alpha_2}, \ldots, v^{\alpha_v}, \ldots\}.$$

Clearly, the equivalence class is a *partition* [21] of twice the number of edges:

$$\sum i\,\alpha_i = 2\,|E|,$$

and α_i, $i = 1,2,\ldots$ is a partition of the number of vertices:

$$\sum \alpha_i = |V|.$$

For trees, the number of edges is one less than the number of vertices; hence, the valence class of a tree satisfies:

$$2 + \sum (i - 2)\alpha_i = 0.$$

(Note that 2-valent vertices do not contribute in this relation.) For trees with n vertices, the valence class $\{1^2, 2^{(n-2)}\}$ belongs to the linear *chain* of vertices. The

valence class $\{1^{(n-1)}, (n-1)\}$ belongs to the "*star*" with $n-1$ branches. Other valence classes often contain two or more non-isomorphic graphs.

4. Lexicographical ordering of valence classes

Graphs can be classified and partially ordered by their valence classes. Randić and Ruch adopt Muirhead's partial sums method [22] for ordering valence classes. We have chosen lexicographical ordering because (i) it is a well ordered relation on the valence classes, and (ii) it accounts for important features of the embedding relation.

In lexicographical order, two valence classes are compared term by term until a *difference* is found. There are two choices: The comparison may begin with the small valences and ascend to the larger ones (a choice we call "outside-to-inside" because the leaves are compared first, and the highest valence vertices last). Or the comparison may begin with the large valences and descend to the smaller ones ("inside-to-outside"). We adopt the outside-to-inside ordering defined as follows; let

$$\alpha = \{1^{\alpha_1}, 2^{\alpha_2}, \ldots\}$$

and

$$\beta = \{1^{\beta_1}, 2^{\beta_2}, \ldots\}.$$

Then, $\alpha < \beta$ (we say α precedes β) if and only if there exists an integer k such that $\alpha_i = \beta_i$ for all $i < k$, and $\alpha_k > \beta_k$ (regardless of the sizes of α_i and β_i for $k < i$). For example, in the outside-to-inside lexicographical ordering: $\{1^6,3,5\} < \{1^6,4^2\} < \{1^5,2^2,5\} < \{1^5,2,3,4\}$ etc. In the inside-to-outside ordering, the same valence classes lie in the order: $\{1^5,2,3,4\} < \{1^6,4^2\} < \{1^5,2^2,5\} < \{1^6,3,5\}$. The two valence classes $\{1^5,2^2,5\}$ and $\{1^6,4^2\}$ are not comparable by the partial sums ordering rule [22]. In our ordering, the stars constitute the smallest valence class and the chains the largest valence class among trees with the same number of vertices.

The lexicographical order of valence classes induces a partial ordering of graphs. If Γ and Γ' have the same valence class, we call them valence class equivalent. The number of non-isomorphic graphs in the valence class α is denoted by $n(\alpha)$. These numbers are easily read from our tabulated results.

As will become evident below, outside-to-inside lexicographical ordering imparts a strong diagonal structure to the subtree embedding relation. We make this structure precise in the theorem proved below; roughly speaking, we show that lexicographical ordering is preserved upon pruning leaves from a tree.

5. Embedded valence classes

We say that the valence class α is embedded in the valence class β if there exists a pair of graphs Γ_1 and Γ_2 having these valence classes and such that $\Gamma_1 \subseteq \Gamma_2$.

For example, if a single leaf be pruned from all graphs in a valence class α in all possible ways, the resulting graphs span several valence classes of one lower degree. For chains, there is only a single valence class in the once pruned subgraphs:

$$1^2\, 2^{n-2} \rightarrow 1^2\, 2^{n-3}.$$

Likewise, for stars:

$$1^{n-1}\,(n-1) \rightarrow 1^{n-2}\,(n-2).$$

These are algebraic expressions for the invariance of the chains or the stars under pruning. Conversely, trees can grow by budding a new leaf on any vertex. The collection of all trees produced when a single bud sprouts from all vertices of a valence class β in all possible ways spans those valence classes α which yield β by pruning.

When valence classes are lexicographically ordered, the collection of embedded valence classes for once pruned trees of the valence class α contains a least and a greatest member. Let us write the valence class of a general tree:

$$\alpha = \{1^{\alpha_1}, i^{\alpha_i}, \ldots\},$$

where i is the smallest non-leaf valence class appearing. Then the valence classes of the collection of once pruned subgraphs are written

$$\alpha - 1 \rightarrow \alpha_<, \ldots, \alpha_{><}, \ldots, \alpha_> .$$

It is then clear that the least, the greatest, and a general intermediate pruned valence class can be written:

$$\alpha_< = \left\{1^{\alpha_1 - 1}, (i-1)^1, i^{\alpha_i - 1}, \ldots, n^{\alpha_n}\right\}$$

$$\alpha_{><} = \left\{1^{\alpha_1 - 1}, i^{\alpha_i}, \ldots, (j-1)^{\alpha_{j-1}+1}, j^{\alpha_j - 1}, \ldots, n^{\alpha_n}\right\}$$

$$\alpha_> = \left\{1^{\alpha_1 - 1}, i^{\alpha_i}, \ldots, j^{\alpha_j}, \ldots, (n-1)^{\alpha_{n-1}+1}, n^{\alpha_n - 1}\right\}.$$

We use i for the smallest valence (greater than 1) appearing in α.

6. The enumeration problem

Suppose Γ is an n-vertex graph and γ is an m-vertex graph. Then the embedding frequency $N(\Gamma, \gamma)$ for γ in Γ is defined to be the number of different subgraphs of Γ isomorphic to γ. Two subgraphs are "different" if their vertex sets

are not the same. When Γ and γ are regarded as row and column labels, respectively, then the array of $N(\Gamma, \gamma)$ is called the embedding frequency matrix. Various classes of graphs are closed under the subgraph embedding relation and correspond to submatrices of N. For example, we limit our discussion to one such class, the trees.

The subgraph enumeration problem is to count the number of times γ is embedded in Γ. A partial solution, in the form of a table of values $N(\Gamma, \gamma)$ through the 201 trees with 10 or fewer vertices, in given in table 2. Tables for larger trees quickly become too large for a journal article (there are 987 trees with 12 or fewer vertices). Therefore, other forms of the solution, such as recursion formulas or generating functions, are more practical as well as more satisfying. Such solutions remain an outstanding unsolved problem.

It is clear that graphs can be partially ordered by the number of vertices. That is, the rows and columns of N are taken in the order: $|\Gamma| = 1, 2, 3, \ldots$. This results in a lower triangular embedding frequency matrix with $N(\Gamma, \Gamma) = 1$ on the diagonal.

Additional approximate triangular structure is induced by ordering the trees according to valence class. That is, within a given $|\Gamma| = n$, we order the graphs by lexicographical order of their valence class. If two graphs have the same valence class, they may be further ordered by the valence classes of their once pruned subgraphs (defined below).

7. Computer program

Computer programs for generating and manipulating graphs have been reported [17,18,23]. Our own computer program to enumerate the subgraphs exploits published algorithms as well as some original features. First, a standard list of all graphs (trees) with 10 or fewer vertices is given. The form of this list is described below. Second, an n-vertex graph Γ is selected; a leaf is removed from Γ to produce an $(n-1)$-vertex graph γ. Third, a search is made through the standard list until γ is found. Lastly, the counter for $N(\Gamma, \gamma)$ is increased by 1. The pruning process is repeated, removing all possible leaves, all possible pairs of vertices, all possible triples, etc. until all connected subgraphs have been generated and counted. The program contains a bootstrap feature: If the given standard list is incomplete, so that γ is not found in step three, then γ is inserted into the list. This feature was used to help to construct the complete standard list of codes for trees.

8. Read's walk around valence codes

Read has described criteria for selecting codes to represent trees [17,23]. We have selected Read's walk around valence code ([17], p. 179), a string of n integers, one for each vertex. In this way, each tree has a unique code, occupying very little

Table 1

Standard list of trees with no more than ten vertices. The first column lists the sequence number of the tree. The second entry in each row is the number of vertices in that graph. The following entries are the Read walk around valence code for the tree as explained in the text

1	1	0								
2	2	1	0							
3	3	2	0	0						
4	4	2	1	0	0					
5	4	3	0	0	0					
6	5	2	1	0	1	0				
7	5	2	2	0	0	0				
8	5	4	0	0	0	0				
9	6	2	1	1	0	1	0			
10	6	2	1	0	2	0	0			
11	6	3	0	1	0	1	0			
12	6	3	2	0	0	0	0			
13	6	2	3	0	0	0	0			
14	6	5	0	0	0	0	0			
15	7	2	1	1	0	1	1	0		
16	7	2	1	2	0	0	1	0		
17	7	2	2	0	1	0	1	0		
18	7	3	1	0	1	0	1	0		
19	7	2	2	0	0	2	0	0		
20	7	3	0	1	0	2	0	0		
21	7	2	1	0	3	0	0	0		
22	7	4	0	0	1	0	1	0		
23	7	3	3	0	0	0	0	0		
24	7	2	4	0	0	0	0	0		
25	7	6	0	0	0	0	0	0		
26	8	2	1	1	1	0	1	1	0	
27	8	2	1	1	0	1	2	0	0	
28	8	2	1	1	0	2	0	1	0	
29	8	3	0	1	1	0	1	1	0	
30	8	2	2	1	0	1	0	1	0	
31	8	2	1	2	0	0	2	0	0	
32	8	3	1	0	1	0	2	0	0	
33	8	3	2	0	1	0	0	1	0	
34	8	2	2	0	1	0	2	0	0	
35	8	2	2	0	2	0	0	1	0	
36	8	2	1	3	0	0	0	1	0	
37	8	2	3	0	0	1	0	1	0	
38	8	4	0	1	0	1	0	1	0	
39	8	3	0	2	0	0	2	0	0	
40	8	2	2	0	0	3	0	0	0	
41	8	3	0	1	0	3	0	0	0	
42	8	4	0	0	1	0	2	0	0	
43	8	2	1	0	4	0	0	0	0	
44	8	5	0	0	0	1	0	1	0	
45	8	4	3	0	0	0	0	0	0	
46	8	3	4	0	0	0	0	0	0	
47	8	2	5	0	0	0	0	0	0	
48	8	7	0	0	0	0	0	0	0	
49	9	2	1	1	1	0	1	1	1	0
50	9	2	1	1	2	0	0	1	1	0
51	9	2	1	2	0	1	0	1	1	0
52	9	2	2	0	1	1	0	1	1	0
53	9	2	1	1	0	2	1	0	1	0
54	9	3	1	0	1	1	0	1	1	0
55	9	2	1	2	0	0	1	2	0	0
56	9	2	2	1	0	2	0	0	1	0
57	9	3	0	1	1	0	1	2	0	0
58	9	3	0	1	1	0	2	0	1	0
59	9	3	2	1	0	1	0	0	1	0
60	9	2	1	2	0	0	2	0	1	0
61	9	2	2	0	1	0	2	0	1	0
62	9	2	2	1	0	1	0	2	0	0
63	9	2	1	1	0	2	0	2	0	0
64	9	2	1	1	0	1	3	0	0	0
65	9	2	1	1	0	3	0	0	1	0
66	9	2	3	0	1	0	1	0	1	0
67	9	4	0	0	1	1	0	1	1	0
68	9	4	1	0	1	0	1	0	1	0
69	9	2	2	0	2	0	0	2	0	0
70	9	3	2	0	2	0	0	0	1	0
71	9	3	1	0	2	0	0	2	0	0
72	9	2	1	3	0	0	0	2	0	0
73	9	2	3	0	0	2	0	0	1	0
74	9	3	1	0	1	0	3	0	0	0
75	9	3	1	3	0	0	0	0	1	0
76	9	3	3	0	0	1	0	0	1	0
77	9	4	0	1	0	1	0	2	0	0
78	9	2	2	0	3	0	0	0	1	0
79	9	2	3	0	0	1	0	2	0	0
80	9	2	1	4	0	0	0	0	1	0
81	9	2	4	0	0	0	1	0	1	0
82	9	5	0	0	1	0	1	0	1	0
83	9	4	0	0	2	0	0	2	0	0
84	9	3	0	2	0	0	3	0	0	0
85	9	2	3	0	0	0	3	0	0	0
86	9	4	0	0	1	0	3	0	0	0
87	9	2	2	0	0	4	0	0	0	0
88	9	3	0	1	0	4	0	0	0	0
89	9	5	0	0	0	1	0	2	0	0
90	9	2	1	0	5	0	0	0	0	0
91	9	6	0	0	0	0	1	0	1	0
92	9	4	4	0	0	0	0	0	0	0
93	9	3	5	0	0	0	0	0	0	0
94	9	2	6	0	0	0	0	0	0	0
95	9	8	0	0	0	0	0	0	0	0
96	10	2	1	1	1	1	0	1	1	1	0
97	10	2	1	1	1	0	1	1	2	0	0
98	10	2	1	1	1	0	1	2	0	1	0
99	10	2	1	1	1	0	2	0	1	1	0
100	10	3	0	1	1	1	0	1	1	1	0

Table 1 (continued)

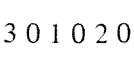
101	10	2	2	1	0	1	1	0	1	1	0
102	10	2	1	2	1	0	1	0	1	1	0
103	10	3	1	1	0	1	1	0	1	1	0
104	10	2	1	1	2	0	0	1	2	0	0
105	10	2	1	2	0	2	0	0	1	1	0
106	10	2	2	0	1	0	2	1	0	1	0
107	10	2	2	0	1	1	0	1	2	0	0
108	10	2	2	0	1	1	0	2	0	1	0
109	10	2	2	0	1	2	0	0	1	1	0
110	10	2	2	0	2	0	1	0	1	1	0
111	10	3	0	1	1	0	2	1	0	1	0
112	10	3	1	0	1	1	0	1	2	0	0
113	10	3	1	0	1	1	0	2	0	1	0
114	10	3	1	1	0	1	1	0	2	0	0
115	10	3	2	0	1	1	0	0	1	1	0
116	10	2	1	2	0	0	2	1	0	1	0
117	10	2	1	2	0	1	0	1	2	0	0
118	10	2	1	2	0	1	0	2	0	1	0
119	10	2	1	1	0	2	1	0	2	0	0
120	10	3	2	1	0	1	0	1	0	1	0
121	10	2	1	1	0	3	0	1	0	1	0
122	10	2	1	1	3	0	0	0	1	1	0
123	10	2	1	3	0	0	1	0	1	1	0
124	10	2	3	0	0	1	1	0	1	1	0
125	10	4	0	1	0	1	1	0	1	1	0
126	10	2	3	1	0	1	0	1	0	1	0
127	10	3	0	1	2	0	0	1	2	0	0
128	10	2	1	2	0	0	2	0	2	0	0
129	10	2	2	2	0	0	2	0	0	1	0
130	10	3	2	1	0	2	0	0	0	1	0
131	10	3	0	1	2	0	0	2	0	1	0
132	10	3	0	2	0	1	0	2	0	1	0
133	10	3	2	1	0	1	0	0	2	0	0
134	10	3	0	1	1	0	2	0	2	0	0
135	10	2	2	0	1	0	2	0	2	0	0
136	10	2	2	1	0	2	0	0	2	0	0
137	10	2	1	2	0	0	1	3	0	0	0
138	10	2	2	1	0	3	0	0	0	1	0
139	10	2	3	0	1	0	1	0	2	0	0
140	10	2	3	0	1	0	2	0	0	1	0
141	10	3	0	1	1	0	1	3	0	0	0
142	10	3	0	1	1	0	3	0	0	1	0
143	10	3	3	0	0	1	0	1	0	1	0
144	10	3	3	0	1	0	1	0	0	1	0
145	10	4	0	0	1	1	0	1	2	0	0
146	10	4	0	0	1	1	0	2	0	1	0
147	10	2	1	2	0	0	3	0	0	1	0
148	10	2	2	0	1	0	1	3	0	0	0
149	10	2	2	0	1	0	3	0	0	1	0
150	10	2	2	1	0	1	0	3	0	0	0
151	10	2	1	1	0	2	0	3	0	0	0
152	10	2	1	1	0	3	0	0	2	0	0
153	10	4	1	0	1	0	1	0	2	0	0
154	10	2	1	1	0	1	4	0	0	0	0
155	10	2	1	1	0	4	0	0	0	1	0
156	10	2	4	0	0	1	0	1	0	1	0
157	10	5	0	0	0	1	1	0	1	1	0
158	10	5	0	1	0	1	0	1	0	1	0
159	10	3	2	0	0	2	0	0	2	0	0
160	10	3	2	0	2	0	0	0	2	0	0
161	10	2	2	0	2	0	0	3	0	0	0
162	10	2	2	0	3	0	0	0	2	0	0
163	10	2	3	0	0	2	0	0	2	0	0
164	10	3	2	0	3	0	0	0	0	1	0
165	10	3	3	0	0	1	0	0	2	0	0
166	10	3	3	0	0	2	0	0	0	1	0
167	10	4	0	1	0	2	0	0	2	0	0
168	10	3	1	0	2	0	0	3	0	0	0
169	10	2	1	3	0	0	0	3	0	0	0
170	10	4	3	0	0	1	0	0	0	1	0
171	10	2	3	0	0	1	0	3	0	0	0
172	10	2	3	0	0	3	0	0	0	1	0
173	10	4	0	1	0	1	0	3	0	0	0
174	10	2	1	4	0	0	0	0	2	0	0
175	10	2	4	0	0	0	2	0	0	1	0
176	10	3	1	0	1	0	4	0	0	0	0
177	10	3	1	4	0	0	0	0	0	1	0
178	10	3	4	0	0	0	1	0	0	1	0
179	10	5	0	0	1	0	1	0	2	0	0
180	10	2	2	0	4	0	0	0	0	1	0
181	10	2	4	0	0	0	1	0	2	0	0
182	10	2	1	5	0	0	0	0	0	1	0
183	10	2	5	0	0	0	0	1	0	1	0
184	10	6	0	0	0	1	0	1	0	1	0
185	10	3	0	3	0	0	0	3	0	0	0
186	10	4	0	0	2	0	0	3	0	0	0
187	10	5	0	0	0	2	0	0	2	0	0
188	10	3	0	2	0	0	4	0	0	0	0
189	10	2	3	0	0	0	4	0	0	0	0
190	10	5	0	0	0	1	0	3	0	0	0
191	10	4	0	0	1	0	4	0	0	0	0
192	10	2	2	0	0	5	0	0	0	0	0
193	10	3	0	1	0	5	0	0	0	0	0
194	10	6	0	0	0	0	1	0	2	0	0
195	10	2	1	0	6	0	0	0	0	0	0
196	10	7	0	0	0	0	0	1	0	1	0
197	10	5	4	0	0	0	0	0	0	0	0
198	10	4	5	0	0	0	0	0	0	0	0
199	10	3	6	0	0	0	0	0	0	0	0
200	10	2	7	0	0	0	0	0	0	0	0
201	10	9	0	0	0	0	0	0	0	0	0

storage space, which becomes its name in a list used for table look-up. The greatest shortcomings of these codes is that they apply only to trees and not to graphs with cycles. Table 1 lists the Read codes for the 201 trees with up to 10 vertices in the order of their valence classes as previously described.

Read codes are also partitions; the sum of integers in the code is one less than the number of vertices. However, the order of integers in the code is important. Nevertheless, it is sometimes convenient to use the same power notation for codes that we have used for valence classes.

Read describes how to decode the walk around valency codes. For example, consider the following code: 3 0 1 0 2 0 0. The first vertex, with valence 3, has three branches. The first branch, starting with the second vertex has no (0) further branches. The second branch, starting with the third vertex, has one (1) further branch on which lies the fourth vertex, which in turn has no (0) further branches. The third branch from vertex 1, starting with vertex 5, has two (2) further branches, each terminating in a leaf (0, 0):

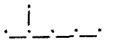
3 0 1 0 2 0 0 →



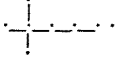
In this way, each code of table 1 can be decoded into its graph.

9. Results

The results of our subtree enumerations are presented in table 2 and depicted graphically in fig. 1. Trees are ordered in these presentations according to table 1, with larger valence classes above smaller ones. The corresponding rows and columns of table 2 and fig. 1 provide the embedding frequencies of these trees. Zeros become more and more prevalent with larger trees. Ratios of numbers of zero entries to total number of zero and non-zero entries in all rows with $|\Gamma| = 4, 5, \ldots, 10$ are 1/9, 5/21, 26/69, 114/220, 551/851, 2550/3384, and 13181/15741, respectively. To conserve space in table 2, we take advantage of the sparsity of the frequency matrix (81% of the 20,301 entries in table 2 are zero). For example, the 25th row is 7, 6, 15, 0, 20, 0, 0, 15, 0, 0, 0, 0, 0, 6, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1. We abbreviate this by indicating the length of a string of repeated zeros in parenthesis; thus, row 25 is written: 7, 6, 5, 0, 20, (2), 15, (5), 6, (10), 1. This makes the matrix more difficult to read, but reduces the longest row from 201 columns to only 42 columns. For example, the 10th tree



is embedded three times in the 21st tree



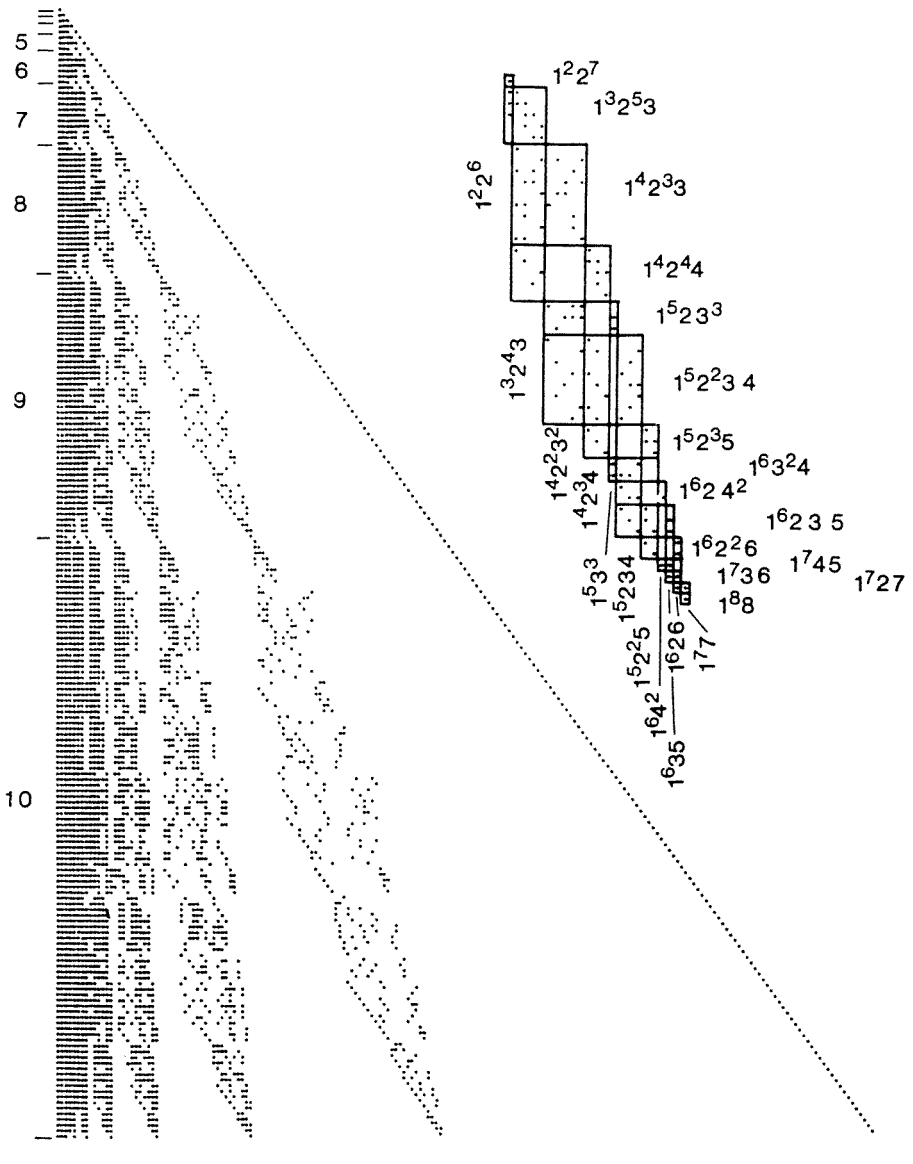


Fig. 1. Zeta function for trees: $\zeta(\Gamma, \gamma) = 1$ if $\gamma \subseteq \Gamma$. Depiction of the non-zero embedding frequencies for trees having no more than ten vertices. The entry in the Γ row and γ column is blank if γ is not embedded in Γ and contains a dot otherwise. Trees are ordered as indicated in table 1. The inset shows, in a larger scale, the block of 8-vertex trees embedded in 9-vertex trees; also shown are the valence classes of these trees and their once pruned subtrees to illustrate the ordering theorem.

Table 2

Embedding frequency matrix for trees with no more than ten vertices. The first column is the sequence number of the graph Γ (table 1). Columns are ordered the same as the rows. An entry $N(\Gamma, \gamma)$ is the number of times γ appears as a subgraph within Γ. Strings of zeros are abbreviated to conserve space: a series of 10 zeros is written (10)

1	1																
2	2	1															
3	3	2	1														
4	4	3	2	1													
5	4	3	3	0	1												
6	5	4	3	2	0	1											
7	5	4	4	2	1	0	1										
8	5	4	6	0	4	(2)	1										
9	6	5	4	3	0	2	(2)	1									
10	6	5	5	3	1	2	1	(2)	1								
11	6	5	5	4	1	1	2	(3)	1								
12	6	5	6	4	2	0	4	(4)	1								
13	6	5	7	3	4	0	3	1	(4)	1							
14	6	5	10	0	10	(2)	5	(5)	1								
15	7	6	5	4	0	3	(2)	2	(5)	1							
16	7	6	6	4	1	3	1	0	2	1	(5)	1					
17	7	6	6	5	1	3	2	0	1	1	1	(5)	1				
18	7	6	6	6	1	3	3	(3)	3	(6)	1						
19	7	6	7	4	2	4	2	(2)	4	(8)	1						
20	7	6	7	6	2	2	5	(2)	1	2	1	(7)	1				
21	7	6	8	4	4	3	3	1	0	3	(2)	1	(7)	1			
22	7	6	8	6	4	1	6	1	(2)	2	0	2	(8)	1			
23	7	6	9	6	5	0	9	1	(3)	3	2	(9)	1				
24	7	6	11	4	10	0	6	5	(4)	4	1	(9)	1				
25	7	6	15	0	20	(2)	15	(5)	6	(10)	1						
26	8	7	6	5	0	4	(2)	3	(5)	2	(10)	1					
27	8	7	7	5	1	4	1	0	3	1	(4)	2	1	(10)	1		
28	8	7	7	6	1	4	2	0	3	1	1	(3)	1	1	1	(10)	1
29	8	7	7	6	1	5	2	0	2	2	1	(3)	1	0	2	(11)	1
30	8	7	7	7	1	5	3	0	2	1	3	(5)	2	1	(11)	1	

31	8	7	8	5	2	4	2	0	4	2	(5)	4	(14)	1													
32	8	7	8	8	2	5	6	(2)	2	5	1	(5)	2	0	2	(11)	1										
33	8	7	8	8	2	4	6	0	1	2	4	1	(4)	2	(2)	2	(12)	1									
34	8	7	8	6	2	5	3	0	2	4	1	(4)	1	2	0	1	(14)	1									
35	8	7	8	7	2	4	5	0	2	2	2	1	(3)	1	2	(2)	1	(14)	1								
36	8	7	9	5	4	4	3	1	3	3	(2)	1	(2)	3	(4)	1	(14)	1									
37	8	7	9	7	4	4	6	1	1	3	2	0	2	(3)	2	(3)	1	1	(14)	1							
38	8	7	9	9	4	3	9	1	(2)	6	0	3	(4)	1	(3)	3	(15)	1									
39	8	7	9	8	3	4	8	(2)	4	4	2	(6)	1	4	(18)	1											
40	8	7	10	5	5	6	4	1	0	9	(2)	1	(5)	3	0	2	(18)	1									
41	8	7	10	8	5	3	10	1	0	3	3	3	2	(6)	3	1	0	1	(17)	1							
42	8	7	10	9	5	2	12	1	0	1	4	3	3	(6)	2	0	2	1	(18)	1							
43	8	7	12	5	10	4	6	5	0	6	(2)	4	1	(6)	4	(2)	1	(18)	1								
44	8	7	12	8	10	1	12	5	(2)	3	0	8	1	(7)	3	0	2	(19)	1								
45	8	7	12	9	8	0	18	2	(3)	9	6	(9)	6	(21)	1												
46	8	7	13	8	11	0	16	5	(3)	6	8	1	(8)	4	2	(21)	1										
47	8	7	16	5	20	0	10	15	(4)	10	6	(9)	5	1	(21)	1											
48	8	7	21	0	35	(2)	35	(5)	21	(10)	7	(22)	1														
49	9	8	7	6	0	5	(2)	4	(5)	3	(10)	2	(22)	1													
50	9	8	8	6	1	5	1	0	4	1	(4)	3	1	(9)	2	1	(22)	1									
51	9	8	8	7	1	5	2	0	4	1	1	(3)	3	1	1	(8)	1	1	1	(22)	1						
52	9	8	8	7	1	6	2	0	4	2	1	(3)	2	1	2	(8)	1	0	1	1	(22)	1					
53	9	8	8	8	1	6	3	0	4	1	3	(3)	2	1	2	1	(9)	2	0	1	(22)	1					
54	9	8	8	8	1	7	3	0	4	2	3	(3)	1	0	4	1	(10)	1	2	(23)	1						
55	9	8	9	6	2	5	2	0	4	2	(4)	4	2	(10)	4	(27)	1										
56	9	8	9	9	2	7	6	0	3	3	5	1	(3)	1	3	2	0	2	(9)	2	0	1	(2)	1	(20)	1	
57	9	8	9	7	2	7	3	0	3	5	1	(3)	2	1	3	0	1	(7)	1	0	2	(4)	1	(22)	1		
58	9	8	9	9	2	6	6	0	3	3	4	1	(2)	1	1	4	(2)	2	(7)	1	1	(3)	1	0	1	(22)	1
59	9	8	9	10	2	7	7	0	2	3	7	1	(4)	4	2	0	3	(9)	1	1	2	(25)	1				
60	9	8	9	7	2	5	3	0	5	2	1	(3)	2	4	1	(9)	1	2	(2)	1	(28)	1					
61	9	8	9	8	2	6	4	0	4	4	2	(3)	1	2	4	0	1	(8)	2	(5)	2	(26)	1				
62	9	8	9	8	2	7	4	0	4	4	3	(4)	2	4	1	1	(10)	2	(3)	2	(27)	1					
63	9	8	9	8	2	5	5	0	4	2	2	1	(2)	2	2	(2)	1	(6)	1	1	2	(6)	1	(27)	1		
64	9	8	10	6	4	5	3	1	4	3	(2)	1	0	3	3	(4)	1	(5)	3	(8)	1	(27)	1				
65	9	8	10	8	4	5	6	1	4	3	2	0	2	0	1	3	2	(3)	1	1	(5)	2	(7)	1	1	(27)	1
66	9	8	10	10	4	6	9	1	2	3	6	0	3	(3)	4	1	(2)	1	3	(7)	1	(6)	2	1	(27)	1	

Table 2 (continued)

67	9	8	10	8	4	7	6	1	2	6	2	0	2	0	1	0	4	(3)	2	1	(6)	2	(7)	2	(29)	1			
68	9	8	10	12	4	6	12	1	(2)	12	0	4	(4)	4	(3)	6	(15)	4	(29)	1									
69	9	8	10	8	3	6	6	0	4	5	2	1	(3)	4	4	0	1	1	(10)	1	(2)	2	2	(33)	1				
70	9	8	10	10	3	6	9	0	2	5	6	2	(3)	1	4	0	1	5	(12)	2	1	1	(3)	1	(30)	1			
71	9	8	10	10	3	8	9	(2)	6	8	2	(5)	4	1	6	(11)	4	(6)	1	(31)	1								
72	9	8	11	6	5	5	4	1	6	4	(2)	1	(2)	9	(4)	1	(9)	3	(4)	2	(35)	1							
73	9	8	11	10	5	5	12	1	2	4	4	3	3	(2)	1	4	(2)	2	1	2	1	(11)	2	0	2	(4)	1	(30)	1
74	9	8	11	10	5	7	11	1	0	6	7	3	2	(4)	3	0	6	2	0	1	(8)	3	(8)	2	(32)	1			
75	9	8	11	7	5	7	5	1	3	9	1	0	1	(2)	3	3	0	3	0	2	(12)	3	0	1	(3)	1	(34)	1	
76	9	8	11	11	5	5	13	1	1	4	7	3	3	(3)	3	(2)	5	1	2	1	(9)	2	(3)	1	(3)	1	1	(33)	1
77	9	8	11	12	5	5	15	1	0	2	10	3	4	(4)	2	0	4	0	5	1	(8)	1	(5)	2	(3)	2	(34)	1	
78	9	8	11	9	5	5	10	1	3	4	3	3	2	(2)	3	3	(2)	3	1	0	1	(11)	3	1	(4)	1	(36)	1	
79	9	8	11	8	5	7	7	1	2	9	2	0	2	(2)	1	4	0	3	0	2	1	(11)	2	(2)	2	(2)	1	(38)	1
80	9	8	13	6	10	5	6	5	4	6	(2)	4	1	0	6	(4)	4	(2)	1	(11)	4	(6)	1	(36)	1				
81	9	8	13	9	10	5	12	5	1	6	3	0	8	1	(2)	3	(3)	4	3	0	2	(12)	3	(5)	1	1	(36)	1	
82	9	8	13	12	10	3	18	5	(2)	9	0	12	1	(3)	1	(3)	9	0	3	(13)	2	(5)	3	(37)	1				
83	9	8	12	12	6	4	18	1	0	4	8	6	4	(5)	1	8	0	4	2	(15)	2	(2)	4	(40)	1				
84	9	8	12	10	6	6	13	1	0	9	6	4	2	(5)	3	9	2	0	1	(15)	3	1	2	(42)	1				
85	9	8	13	6	8	9	6	2	0	18	(2)	2	(5)	9	0	6	(18)	6	(44)	1									
86	9	8	13	12	8	3	21	2	0	3	6	9	7	(6)	6	1	3	6	(17)	2	3	(2)	1	(40)	1				
87	9	8	14	6	11	8	7	5	0	16	(2)	4	1	(4)	6	0	8	(2)	1	(15)	4	(2)	2	(43)	1				
88	9	8	14	10	11	4	17	5	0	6	4	6	8	1	(5)	6	4	0	4	2	(16)	4	0	1	(2)	1	(41)	1	
89	9	8	14	12	11	2	22	5	0	1	6	6	12	1	(5)	3	0	6	4	3	(17)	3	0	2	0	1	(42)	1	
90	9	8	17	6	20	5	10	15	0	10	(2)	10	6	(6)	10	(2)	5	1	(17)	5	(3)	1	(42)	1					
91	9	8	17	10	20	1	10	15	(2)	4	0	20	6	(7)	6	0	10	1	(18)	4	(2)	2	(43)	1					
92	9	8	16	12	14	0	30	6	(3)	18	16	1	(8)	18	3	(20)	4	3	(45)	1									
93	9	8	18	10	21	0	25	15	(3)	10	20	6	(8)	10	10	1	(20)	5	2	(45)	1								
94	9	8	22	6	35	0	15	35	(4)	20	21	(9)	15	7	(21)	6	1	(45)	1										
95	9	8	28	0	56	(2)	70	(5)	56	(10)	28	(22)	8	(46)	1														
96	10	9	8	7	0	6	(2)	5	(5)	4	(10)	3	(22)	2	(46)	1													
97	10	9	9	7	1	6	1	0	5	1	(4)	4	1	(9)	3	1	(21)	2	1	(46)	1								
98	10	9	9	8	1	6	2	0	5	1	1	(3)	4	1	1	(8)	3	1	1	(20)	1	1	1	(46)	1				
99	10	9	9	8	1	7	2	0	5	2	1	(3)	4	1	2	(8)	2	1	1	1	(19)	1	0	1	1	(46)	1		
100	10	9	9	8	1	7	2	0	6	2	1	(3)	3	2	2	(8)	2	0	2	1	(19)	1	(2)	2	(47)	1			

Table 2 (continued)

141	10	9	12	8	5	9	5	1	4	10	1	0	1	0	3	3	4	0	3	0	2	(5)	3	0	3	(4)	3	0	1	(3)	1	(16)	3	(6)	1	(10)	1	(65)	1
142	10	9	12	12	5	7	13	1	4	5	7	3	3	0	1	3	6	(2)	5	1	2	1	(4)	2	1	(3)	2	0	3	1	1	(3)	1	1	(15)	2	(6)	1	(10)	1	0	1	(63)	1
143	10	9	12	13	5	9	14	1	2	7	11	3	3	(3)	6	3	0	8	2	2	1	(6)	1	0	3	4	(3)	2	(3)	2	1	(16)	2	(14)	1	0	2	(66)	1
144	10	9	12	14	5	8	16	1	2	5	13	3	4	(3)	6	2	0	7	1	5	1	(6)	1	0	1	4	(3)	2	3	(2)	1	2	(16)	1	(6)	1	(9)	2	1	(66)	1
145	10	9	12	9	5	10	7	1	3	12	2	0	2	0	2	1	6	0	3	0	3	1	(4)	1	0	4	(4)	2	(2)	3	(2)	1	(16)	2	(9)	2	(11)	1	(65)	1
146	10	9	12	12	5	8	13	1	3	7	7	3	3	0	1	1	7	(2)	5	2	2	1	(4)	1	2	(3)	2	0	2	0	3	(3)	1	1	(15)	2	(8)	1	(5)	1	(2)	1	(69)	1
147	10	9	12	9	5	6	7	1	7	4	2	0	2	0	2	9	2	(3)	1	1	(4)	1	4	(2)	3	(4)	2	1	(22)	2	(4)	2	(6)	1	(74)	1
148	10	9	12	8	5	6	5	1	7	4	1	0	1	0	3	9	1	(3)	1	(5)	3	3	(2)	3	(4)	2	(23)	3	(3)	1	(7)	1	(75)	1
149	10	9	12	10	5	8	8	1	5	9	3	0	2	0	1	4	7	0	3	0	2	1	(5)	3	(5)	5	0	1	2	(2)	1	(20)	2	(3)	1	(9)	1	(3)	1	(69)	1
150	10	9	12	9	5	9	6	1	6	9	3	0	1	(2)	6	6	1	3	0	2	(8)	3	(3)	6	0	2	(3)	1	(21)	3	(12)	2	(74)	1
151	10	9	12	10	5	6	10	1	5	4	3	3	2	0	3	4	3	(2)	3	1	0	1	(3)	3	3	(6)	3	1	(4)	1	(21)	3	1	(13)	1	(72)	1
152	10	9	12	11	5	6	12	1	5	4	4	3	3	0	2	4	4	(2)	2	1	2	1	(3)	1	4	(6)	2	1	2	(4)	1	(20)	2	0	2	(7)	1	(78)	1
153	10	9	12	15	5	9	18	1	0	3	18	3	5	(4)	7	0	6	0	9	1	(8)	3	(5)	7	(3)	3	(25)	2	(8)	3	(75)	1
154	10	9	14	7	10	6	6	5	5	6	(2)	4	1	4	6	(4)	4	(2)	1	(2)	6	(8)	4	(6)	1	(20)	4	(15)	1	(73)	1
155	10	9	14	10	10	6	12	5	5	6	3	0	8	1	1	6	3	(3)	4	3	0	2	(3)	3	(7)	4	3	(5)	1	1	(20)	3	(14)	1	1	(73)	1
156	10	9	14	13	10	7	18	5	2	6	9	0	12	1	(2)	6	1	(2)	4	9	0	3	(5)	1	(6)	6	2	(4)	1	3	(21)	2	(14)	2	1	(73)	1
157	10	9	14	10	10	9	12	5	2	12	3	0	8	1	1	0	6	(3)	8	3	0	2	(4)	3	(7)	6	(5)	2	1	(22)	3	(13)	2	(75)	1
158	10	9	14	16	10	6	24	5	(2)	18	0	16	1	(3)	4	(3)	18	0	4	(13)	8	(5)	6	(23)	1	(13)	4	(75)	1
159	10	9	12	12	4	12	12	(2)	12	12	3	(5)	8	3	12	(11)	12	(6)	3	(31)	6	(87)	1
160	10	9	12	12	4	8	12	0	4	8	8	3	(3)	4	8	0	2	8	(10)	1	0	4	4	4	(3)	2	(29)	2	4	(89)	1
161	10	9	13	9	6	8	8	1	6	10	2	1	1	(2)	9	6	0	3	1	2	(9)	3	(2)	6	3	2	(3)	1	(28)	3	(2)	1	(2)	2	(85)	1
162	10	9	13	10	6	7	11	1	6	7	3	3	2	(2)	9	6	0	1	3	1	0	1	(7)	3	(2)	3	6	2	(4)	1	(27)	3	(2)	1	(5)	2	(83)	1
163	10	9	13	11	6	8	13	1	4	10	4	3	3	(2)	4	8	0	3	2	2	2	1	(7)	1	(2)	4	4	0	4	(2)	1	0	1	(26)	2	(3)	2	(5)	2	(83)	1
164	10	9	13	12	6	8	14	1	3	10	8	4	2	(2)	3	6	0	3	10	2	0	1	(9)	3	3	3	1	(2)	3	1	2	(28)	3	(4)	1	(2)	1	(5)	1	(79)	1
165	10	9	13	13	6	8	16	1	2	10	10	4	3	(2)	1	6	0	3	11	2	2	1	(9)	4	2	1	0	2	0	3	1	2	1	(27)	2	(5)	2	(2)	1	(4)	1	(80)	1
166	10	9	13	14	6	7	19	1	2	7	11	6	4	(2)	1	6	0	1	11	1	4	2	(9)	4	1	2	0	2	0	2	0	1	4	(27)	2	(2)	1	(2)	2	(6)	1	(82)	1
167	10	9	13	15	6	8	21	1	0	6	16	6	5	(4)	4	1	12	0	8	2	(8)	4	(5)	4	2	(2)	6	(28)	1	(5)	4	(5)	1	(83)	1
168	10	9	13	12	6	11	14	1	0	13	11	4	2	(4)	6	3	13	3	0	1	(8)	9	(6)	3	1	3	(29)	3	(2)	2	(9)	1	(83)	1
169	10	9	14	7	8	6	6	2	9	6	(2)	2	(2)	18	(4)	2	(9)	9	(4)	6	(35)	6	(96)	1
170	10	9	14	15	8	6	24	2	1	6	12	9	8	(3)	4	(2)	12	2	6	6	(9)	4	(3)	2	(3)	4	6	(2)	1	(30)	4	(9)	2	(83)	1
171	10	9	14	9	8	10	9	2	3	18	2	0	3	(2)	3	6	0	9	0	6	1	(11)	6	0	1	3	(2)	6	(34)	2	(3)	3	(5)	1	(85)	1
172	10	9	14	13	8	6	21	2	3	6	6	9	7	(2)	3	6	(2)	6	2	5	6	(11)	6	1	3	(3)	2	3	(2)	1	(27)	3	(4)	2	(7)	1	(85)	1
173	10	9	14	15	8	7	24	2	0	6	14	9	8	(4)	3	0	12	2	7	6	(8)	3	(5)	3	(2)	4	6	(2)	1	(28)	1	(2)	3	(8)	2	(86)	1
174	10	9	15	7	11	6	7	5	8	7	(2)	4	1	0	16	(4)	4	(2)	1	(6)	6	(4)	8	(6)	1	(28)	4	(7)	2	(93)	1
175	10	9	15	13	11	6	22	5	2	7	6	6	12	1	0	1	6	(2)	3	4	6	4	3	(10)	3	0	6	(4)	3	1	2	0	1	(26)	3	(7)	2	(7)	1	(85)	1

176	10	9	15	12	11	9	18	5	0	12	9	6	8	1	(3)	4	0	12	8	0	4	2	(7)	6	(8)	8	0	2	(2)	1	(27)	4	(13)	2	(87)	1
177	10	9	15	8	11	9	8	5	4	16	1	0	4	1	0	6	4	0	6	0	8	(2)	1	(9)	6	0	4	(3)	4	(2)	2	(31)	4	(4)	1	(6)	1	(89)	1
178	10	9	15	14	11	6	23	5	1	7	10	6	12	1	(2)	4	(2)	9	4	6	4	3	(8)	3	(3)	3	(3)	4	3	1	2	0	1	(29)	3	(4)	1	(6)	1	1	(88)	1
179	10	9	15	16	11	5	28	5	0	2	15	6	16	1	(3)	2	0	6	0	15	4	4	(7)	1	(5)	4	(3)	6	0	5	0	1	(30)	2	(4)	2	(6)	2	(89)	1
180	10	9	15	11	11	6	17	5	4	7	4	6	8	1	0	6	4	(2)	6	4	0	4	2	(10)	6	4	(4)	4	0	1	(2)	1	(31)	4	0	1	(7)	1	(91)	1
181	10	9	15	10	11	9	13	5	2	16	3	0	8	1	0	1	6	0	6	0	8	3	0	2	(9)	3	(2)	6	(2)	4	(2)	2	1	(34)	3	0	2	(5)	1	(93)	1
182	10	9	18	7	20	6	10	15	5	10	(2)	10	6	0	10	(4)	10	(2)	5	1	(10)	10	(6)	5	(3)	1	(32)	5	(9)	1	(91)	1
183	10	9	18	11	20	6	20	15	1	10	4	0	20	6	(2)	4	(3)	10	6	0	10	1	(11)	6	(5)	5	4	(2)	2	(33)	4	(8)	1	1	(91)	1
184	10	9	18	15	20	3	30	15	(2)	12	0	30	6	(3)	1	(3)	18	0	15	1	(12)	3	(5)	12	(2)	3	(34)	3	(8)	3	(92)	1
185	10	9	15	12	9	9	18	2	0	18	9	6	4	(5)	9	18	6	0	2	(15)	9	6	6	(42)	6	1	(99)	1
186	10	9	15	15	9	6	27	2	0	9	12	12	8	(5)	3	18	2	6	7	(15)	6	1	4	9	(2)	1	(37)	3	2	0	2	(99)	1
187	10	9	16	16	12	4	32	5	0	4	12	12	16	1	(4)	1	12	0	12	8	4	(14)	3	(2)	12	0	4	0	2	(36)	3	(5)	4	(97)	1
188	10	9	16	12	12	8	20	5	0	16	8	7	8	1	(4)	6	16	8	0	4	2	(14)	6	4	8	0	2	(2)	1	(37)	4	(2)	1	2	(99)	1
189	10	9	17	7	14	12	9	6	0	30	(2)	5	1	(4)	18	0	16	(2)	1	(15)	18	(2)	3	(41)	4	0	3	(101)	1
190	10	9	17	16	14	3	36	6	0	3	9	18	20	1	(5)	9	1	9	18	4	(16)	3	9	0	3	4	3	(39)	3	(2)	3	(2)	1	(97)	1
191	10	9	17	15	14	4	33	6	0	6	8	18	17	1	(5)	12	4	4	18	3	(16)	8	6	1	0	4	3	(39)	4	0	2	(3)	1	(98)	1
192	10	9	19	7	21	10	11	15	0	25	(2)	10	6	(4)	10	0	20	(2)	5	1	(14)	10	(2)	10	(3)	1	(39)	5	(2)	2	(101)	1
193	10	9	19	12	21	5	26	15	0	10	5	10	20	6	(5)	10	10	0	10	10	1	(15)	10	0	5	(2)	5	2	(40)	5	0	1	(2)	1	(99)	1
194	10	9	19	15	21	2	35	15	0	1	8	10	30	6	(5)	4	0	12	10	15	1	(16)	6	0	8	0	5	3	(41)	4	0	2	0	1	(100)	1
195	10	9	23	7	35	6	15	35	0	15	(2)	20	21	(6)	20	(2)	15	7	(17)	15	(3)	6	1	(41)	6	(3)	1	(100)	1
196	10	9	23	12	35	1	30	35	(2)	5	0	40	21	(7)	10	0	30	7	(18)	10	(2)	12	1	(42)	5	(2)	2	(101)	1
197	10	9	20	16	20	0	48	10	(3)	36	32	2	(8)	48	8	(20)	16	12	(45)	8	(104)	1
198	10	9	21	15	24	0	45	16	(3)	30	35	6	(8)	40	15	1	(19)	10	15	3	(44)	5	3	(104)	1
199	10	9	24	12	36	0	36	35	(3)	15	40	21	(8)	20	30	7	(20)	15	12	1	(44)	6	2	(104)	1
200	10	9	29	7	56	0	21	70	(4)	35	56	(9)	35	28	(21)	21	8	(45)	7	1	(104)	1
201	10	9	36	0	84	(2)	126	(5)	126	(10)	84	(22)	36	(46)	9	(105)	1

The data provided here are useful for many purposes. In a later paper, we illustrate several physico-chemical applications. These data have suggested to us certain mathematical properties of the subtree embedding relation. The ordering theorem proved below accounts for the approximate block diagonal structure of fig. 1. Additional properties may be suggested to other readers.

Certain values in the embedding frequency matrix are elementary. For example, the first two columns $N(\Gamma, \ . \ )$ and $N(\Gamma, \ . \_ . \ )$ are merely the numbers of vertices and numbers of edges in Γ. Other values are almost as simple, e.g. embedding frequencies of chains can be counted in the distance matrix (e.g. $N(\Gamma, \ . \_ . \_ . \ )$ is the number of 2's in the distance matrix of Γ). Embedding frequencies of branched subtrees are not so simple; it is their entries in table 2 that carry the most interest.

A subset of graphs is closed under the embedding relation if all their subgraphs are also contained in the subset. A common example is the set of 4-graphs which contains all graphs having 1, 2, 3, or 4-valent vertices. Also, stars and chains are each closed under the embedding relationship. All subgraphs of stars are other stars and those of chains are other chains. (The trees . _ . and . _ . _ . are at once stars as well as chains.)

Let S_n denote a *star* consisting of a central vertex with valence $n-1$ joined to $n-1$ leaves. The Read code of S_n is $(n-1)\,0^{(n-1)}$. Subtrees are formed by successively pruning leaves to produce S_m with code $(m-1)\,0^{(m-1)}$. It is easy to see that the embedding frequencies of stars are given by binomial coefficients

$$N(S_n, S_m) = \binom{n-1}{m-1},$$

except for the first column: $N(S_n, S_1) = n$.

Let C_n denote the chain of n vertices with no side branches. The Read code of C_n is $2\,1^{(n-3)/2}\,0\,1^{(n-3)/2}\,0$ if n is odd; it is $2\,1^{(n-2)/2}\,0\,1^{(n-4)/2}\,0$ if n is even. The only subtrees of a chain are smaller chains and the embedding frequency of the m-chain within the n-chain is

$$N(C_n, C_m) = n - m + 1.$$

10. Pruned trees

Blocks representing single and multiple prunings are conspicuous in the embedding frequencies depicted in fig. 1. The once-pruned blocks, consisting of all rows with n vertices and all columns with $n-1$ vertices are achieved when one leaf is removed in every possible way from all n-vertex trees. Twice-pruned blocks result from removing a second leaf, and so forth. Of course, the non-pruned block is completely diagonal, showing only the elements $N(\Gamma, \Gamma) = 1$. Blocks become more dense as leaves are pruned away from the starting graphs.

Each once-pruned block shows a concentration of subtrees near its diagonal. This trend toward the diagonal is enhanced by ordering according to valence class and may be expressed in the form: subtrees can be less "branched" but cannot be more "branched" than the tree out of which they are pruned. In mathematical terms, the diagonal structure of once-pruned subtrees is required by the ordering theorem. The trend for the diagonal is "remembered" in twice-pruned subtrees, but these blocks are even more dense since pruning leaves also reduces the number of valence classes.

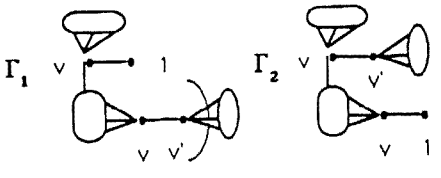
Similarity among once-pruned blocks is noteworthy. The embedding frequencies of once-pruned subgraphs of n-vertex trees have a similar structure to the embedding frequencies of once-pruned subgraphs of $(n-1)$-vertex trees. That is, the n-vertex embeddings "inherit" structure from the embeddings within their subtrees. This property, evident by visual examination of fig. 1, holds promise that embedding frequencies of large graphs can be computed with the aid of embedding frequencies of their subgraphs. We have not found a way to simplify subtree enumeration for n-vertex trees by using their inheritance from $(n-1)$-vertex subtrees.

11. Ordering theorem for embedding frequencies

The structure of the subtree embedding frequency matrix is further explained by the following theorem. The theorem concerns pruning leaves from all graphs in a valence class. For this we need to know there exist graphs in the valence class having at least one leaf on any higher valent vertex. To see this, consider a graph Γ_1 in valence class

$$\alpha = \{1^{\alpha_1}, 2^{\alpha_2}, \ldots, i^{\alpha_i}, \ldots\}.$$

A given vertex, with valence $v > 1$ may, or may not, be joined by an edge to a leaf. We will show that Γ_1 is a valence class equivalent to another graph Γ_2, in which the given vertex *is* joined to a leaf. For this, consider the diagram:



The vertices are labeled by their valences in both graphs Γ_1 and Γ_2. In Γ_1 there is no leaf joined to v; then choose any vertex v' joined to v – this will become a leaf of Γ_2. Now sever all edges joining v' to other vertices except (v, v'). Re-join the severed edges to any leaf of the graph, say the one joined to v''. The result will be Γ_2, which has precisely the same valences as Γ_1 but has a leaf joined to v. This construction

shows that in the valence class α there exists at least one graph with a leaf joined to any of the higher valent vertices that may be present.

For the theorem, consider two valence classes, α and β, and the trees $\{\Gamma_1$, ranging over all $n(\alpha)$ non-isomorphic trees with valence class $\alpha\}$ and $\{\Gamma_2$, ranging over $n(\beta)$ trees with valence class $\beta\}$, $|\Gamma_1| = |\Gamma_2|$. The once-pruned subtrees are $\{\Gamma_1 - 1\}$ and $\{\Gamma_2 - 1\}$, respectively. Of course, the trees of $\{\Gamma_1 - 1\}$ and $\{\Gamma_2 - 1\}$ may be ordered by their valence classes also. The theorem concerns these orders.

THEOREM

If $\alpha < \beta$, then the greatest valence class of $\{\Gamma_1 - 1\}$ does not exceed the greatest valence class of $\{\Gamma_2 - 1\}$ and the least valence class of $\{\Gamma_1 - 1\}$ does not exceed the least valence class of $\{\Gamma_2 - 1\}$.

Proof

Write

$$\alpha = \{1^{\alpha_1}, i^{\alpha_i}, \ldots, j^{\alpha_j}, \ldots\}$$

and

$$\beta = \{1^{\beta_1}, k^{\beta_k}, \ldots, j^{\beta_j}, \ldots .\}$$

By hypothesis $\alpha < \beta$; let j be the smallest valence at which the partitions α and β differ: $\alpha_j > \beta_j$. A leaf can be removed from vertices of each valence and the valence classes of trees pruned from Γ_1 can be listed in lexicographical order:

$$\alpha_< = \left\{1^{\alpha_1 - 1}, (i-1)^1, i^{\alpha_i - 1}, \ldots, n^{\alpha_n}\right\}$$

$$\alpha_{><} = \left\{1^{\alpha_1 - 1}, i^{\alpha_i}, \ldots, (j-1)^{\alpha_{j-1}+1}, j^{\alpha_j - 1}, \ldots, n^{\alpha_n}\right\}$$

$$\alpha_> = \left\{1^{\alpha_1 - 1}, i^{\alpha_i}, \ldots, j^{\alpha_j}, \ldots, (n-1)^{\alpha_{n-1}+1}, n^{\alpha_n - 1}\right\} .$$

Here we use i for the smallest valence in Γ_1 after leaves ($\alpha_t = 0$ for $1 < t < i$) and $i \leqslant j \leqslant n$, with n being the largest valence. Similarly, the valence classes of once-pruned subtrees of Γ_2 are:

$$\beta_< = \left\{1^{\beta_1 - 1}, (k-1)^1, k^{\beta_k - 1}, \ldots, m^{\beta_m}\right\}$$

$$\beta_{><} = \left\{1^{\beta_1 - 1}, k^{\beta_k}, \ldots, (l-1)^{\beta_{l-1}+1}, l^{\beta_l - 1}, \ldots, m^{\beta_m}\right\}$$

$$\beta_> = \left\{1^{\beta_1 - 1}, k^{\beta_k}, \ldots, l^{\beta_l}, \ldots, (m-1)^{\beta_{m-1}+1}, m^{\beta_m - 1}\right\} .$$

First, consider the greatest valence class of once-pruned subtrees. Compare $\alpha_>$ with $\beta_>$. Then $\alpha_>$ precedes $\beta_>$ for the same reason that α precedes β.

Next, consider the smallest valence classes; compare $\alpha_<$ with $\beta_<$. If $i < k$, then we have $\alpha_< < \beta_<$. If $i = k$ and $\beta_i = \alpha_i$, then $\alpha_<$ precedes $\beta_<$ because $\alpha_s > \beta_s$ for some s larger than k as in α and β. Finally, if $i = k$ and $\alpha_i > \beta_i$, then $\alpha_<$ precedes $\beta_<$ because $\alpha_i - 1 > \beta_i - 1$.

12. Discussion of results

The subgraph enumeration problem, a fundamental problem of mathematical graph theory, is extremely difficult although numerical answers can be found easily for restricted examples such as we have done for trees through 10 vertices. As is often the case, these examples, when viewed *en masse,* reveal new structures in the embedding relation and may assist in the search for general solutions.

Mere embedding data without enumeration, known as the Zeta function or the Riemann function [24] and depicted in fig. 1, reveals several structural details of the subtree relation. Embedding matrices (the zeta function or the embedding frequency matrix) are partitioned into blocks of once-, twice-, etc. pruned trees; the blocks repeat similar forms on different scales and different densities. A refinement of the block structure is seen when trees are ordered by their valence classes; this refinement, in turn, is amplified upon enlarging the graphs. The valence class structure of the embedding relation is a manifestation of the partially ordered valence classes, as demonstrated in our ordering theorem.

Our method for enumerating subtrees works by explicitly constructing all subtrees of Γ rather than by considering all smaller trees, many of which are not contained in Γ. Thus, we never compute $N(\Gamma, \gamma) = 0$; zeros only arise by default when γ is not present as a subtree of Γ. However, our method does not utilize information previously acquired about the subtrees of smaller graphs. We have searched without success for a practical recursive method which would build upon previously computed frequencies.

Some embedding frequencies are completely understood: for example, the stars and the chains. Others exhibit ordering relationships which are evident in the N-matrix and which are partly explained by the ordering theorem. A complete understanding of subtree embedding frequencies would be approached by a recursion formula or a generating function, neither of which seems very simple to find.

13. Conclusions

We have computed the subtree embedding frequency matrix through trees with degree $|\Gamma| = 10$. By ordering the trees according to their valence class, the matrix reveals similar blocks which tend to concentrate entries near the diagonal.

We have established an ordering theorem which helps to explain this structure. We report the complete computer-generated embedding frequency table for trees through 10 vertices for the possible benefit of others, who may search for additional relationships of graph embedding. Our results also include the numbers of non-isomorphic trees in a given valence class. In a later paper, we will exhibit some applications of the embedding frequencies to cluster expansions of physical properties.

Embeddings of graphs with cycles can be studied by similar computer methods. Valence classes and partial ordering of embedded valence classes are useful tools for such graphs also. However, repeated pruning of leaves is not the only way to generate subgraphs of graphs with cycles. Therefore, the valence class ordering theorem is less useful than for trees. The form of the Zeta function is also greatly changed, e.g. the unpruned blocks may no longer be diagonal unless the embedding relation is restricted to, say, convex subgraphs.

The most obvious directions for future work are to extend the embedding frequencies to graphs with cycles, to labeled graphs, and to rooted trees.

Acknowledgements

We are grateful to Professor D.J. Klein for his continued interest in this project. We also had helpful discussions with Mr. Ken Davis. Thanks also go to the Academic Computing Service of Washington State University for generously supporting our computations.

References

[1] *Chemical Applications of Topology and Graph Theory*, ed. R.B. King (Elsevier, 1983).
[2] K. Balasubramanian, Chem. Rev. 85(1985)599.
[3] J.W. Essam, J.W. Kennedy, M. Gordon and P. Whittle, J. Chem. Soc. Faraday II, 73(1977) 1289.
[4] D.J. Klein, Int. J. Quant. Chem. S20(1986)153.
[5] N. Trinajstić, D.J. Klein and M. Randić, Int. J. Quant. Chem. S20(1986)699.
[6] D.H. Rouvray, Chem. Soc. Rev. 3(1972)355;
D.H. Rouvray, Endeavour 34(1975)28.
[7] L.B. Kier and L.H. Hall, *Molecular Connectivity in Chemistry and Drug Research* (Academic Press, 1976).
[8] A. Cayley, Amer. J. Math. 4(1881)266;
A. Cayley, Rep. Brit. Assoc. Adv. Sci. 45(1875)257.
[9] G. Pólya, Comptes Rendus Hebdomadaires des Seances de l'Academie des Sciences (Paris) 201(1935)1167.
[10] H.R. Henze and C.M. Blair, J. Amer. Chem. Soc. 56(1934)157.
[11] J.V. Knop, W.R. Müller, K. Szymanski and N. Trinajstić, *Computer Generation of Certain Classes of Molecules*, Association of Chemists and Technologists of Croatia, Sveucilisna naklada Libir, 41000 Zagrab, Yugoslavia (1985).
[12] *Graphical Enumeration*, ed. F. Harary and E.M. Palmer (Academic Press, 1973).

[13] F. Ruskey, SIAM J. Comput. 10(1981)151.
[14] N. Alon, Israel J. Math. 53(1986)97.
[15] M. Gordon and J.W. Kennedy, J. Chem. Soc. Faraday II, 69(1973)484.
[16] D.J. Klein, G.E. Hite and T.G. Schmalz, J. Compt. Chem. 7(1986)443.
[17] *Graph Theory and Computing*, ed. R.C. Read (Academic Press, 1972).
[18] J.V. Knop, W.R. Müller, Z. Jericević and N. Trinajstić, J. Chem. Inf. Comput. Sci. 21 (1981)91.
[19] I. Gutman and M. Randić, Chem. Phys. Lett. 47(1977)15.
[20] E. Ruch, Theor. Chim. Acta 38(1975)167.
[21] W. Lederman, *Introduction to the Theory of Finite Groups* (Oliver and Boyd, 1957) p. 71.12.
[22] R.F. Muirhead, Proc. Edinburgh Math. Soc. 19(1901)36; 21(1903)144; 24(1906)45.
[23] R.C. Read, J. Chem. Inf. Comput. Sci. 23(1983)135.
[24] J.W. Kennedy and M. Gordon, Ann. New York Acad. Sci. 319(1979)331.