# DISTANCE MATRICES OF TREES

by

Karen L. Collins

B. A., Smith College, Northampton, MA 01063

Submitted to the Department of Mathematics in partial fulfillment of the requirements for the degree of Doctor of Philosophy at the Massachusetts Institute of Technology in May, 1986.



Signature of Author____________________________________

MIT, Department of Mathematics

May 19, 1986

Certified by____________________________________

Professor Richard P. Stanley

Thesis Supervisor

Accepted by____________________________________

Professor Nesmith Ankeny

Chair, Departmental Graduate Committee

# DISTANCE MATRICES OF TREES

by
Karen L. Collins

Submitted to the Department of Mathematics in partial fulfillment of the requirements for the degree of Doctor of Philosophy at the Massachusetts Institute of Technology in May, 1986.

ABSTRACT

Let T be a tree. It is well known that the coefficients of the characteristic polynomial (CP) of the adjacency matrix of T can be expressed as the number of matchings of different sizes in T. Graham and Lovasz have shown that a similar result is true for distance matrices, although matchings must be replaced with forests of a certain number of edges. We prove a generalization of their theorem for trees with weighted edges.

Another theorem of Graham, Hoffman and Hosoya says that the determinant of the distance matrix of any graph can be computed by looking only at the 2-connected pieces of the graph. We consider the weighted case and present a possible approach for paths to proving a conjecture of Graham's about the unimodality of the coefficients of the CP of the distance matrix of a tree.

In the final section, we factor the CP of the distance matrix of the full binary tree into factors with degrees less than log n, where n is the number of vertices of T. This factoring generalizes to full k-nary trees. The factors satisfy an easy recursion, and define the CP of the infinite binary tree. The factoring depends on the automorphisms of the binary tree, but we can produce factors for any tree with a non-trivial involution.

This thesis is dedicated

to

**Michael O. Albertson** and **Lucia B. Krompart**

who sustained me throughout my journey.

## ACKNOWLEDGMENTS

I would like to thank

**Mark Hovey** for many discussions and helpful suggestions.

**Richard Stanley** for getting me money, recommending me for conferences, nagging me to write the planar paper, signing my forms, jumping out of an airplane, learning how to juggle, paying for me to come to Caltech, and suggesting the group representation approach to distance matrix polynomials.

**Ron L. Graham** for suggesting that I write my thesis on distance matrices, for having written GL and GHH, for being my mentor in the AT&T Bell Labs Graduate Research Program for Women, for having me at Murray Hill for the last 5 summers, for inviting me to the Japan and China Graph Theory conferences, for showing me how to juggle cross-handed, and for taking me out to breakfast.

**Phil Hanlon** and **Rick Wilson** for being on my thesis committee and reading the first hundred pages of my thesis.

**Richard Brualdi** for asking for a copy of my thesis at my first talk about it.

**Joan Hutchinson** for asking for a copy of my thesis without having heard any talk about it.

**Fan R. K. Chung** for being my co-mentor before the disvestiture, for writing innumerable letters of recommendation, and for going to the University of Pennsylvania.

**AT&T Bell Labs** for their support in their Graduate Research Program for Women.

**Caltech** for use of facilities, office space and word processing.

**MIT** for its financial support, including an RA during 1984-1985 from **Professor Kleitman's** funds.

**K. L. Collins** for her excellent typing and astonishing ability to compose at the computer.

## STANDARD MATHEMATICAL INTRODUCTION

Trees probably don't need an introduction to those of you reading this, nor is it likely that you need to know what they are or why they are interesting. Trees are connected graphs without cycles, and they form one of the most basic classes of examples in graph theory. Trees are 2-colorable, have no sneaky properties that come from allowing graphs to have lots of very closely interconnecting edges, and, in general, are very well behaved. However, in our efforts to understand the wonderful constructs called graphs, we still are so unknowledgeable that we don't even understand trees.

Some efforts have been made to classify trees using polynomials, where, ideally, we have a unique polynomial with degree n for every tree on n vertices. All attempts so far have failed. The particular cases that I am interested in involve the adjacency matrix and the distance matrix of trees. Now the adjacency matrix of a graph on n vertices with vertex set $v_1, v_2, \ldots, v_n$, is an nxn matrix with entries of 0 and 1; the $ij^{th}$ entry is 1 if $v_i$ and $v_j$ and adjacent, and 0 otherwise. The adjacency matrix of a graph characterizes the graph; therefore it was hoped that the characteristic polynomial of the adjacency matrix would characterize trees.

After two trees were discovered that disproved that idea, other aspects of the adjacency matrix were investigated, including its spectral radius (largest magnitude of an eigenvalue) [CSD]. Another attempt led to the consideration of the characteristic polynomial of distance matrices. The distance matrix of a graph on n vertices with vertex set $v_1, v_2, \ldots, v_n$, is an nxn matrix with the $ij^{th}$ entry the distance between $v_i$ and $v_j$. However, a computer search discovered two non-isomorphic trees with identical distance matrix characteristic polynomials. In a recent paper, McKay eliminates most polynomials classically associated with trees as characterizing trees [McK].

Distance matrices remain something of a mystery. While successive powers of adjacency matrices count the number of paths between any two vertices, successive powers of the distance matrix seem to reveal no information. The coefficients of the characteristic polynomial of the adjacency matrix count matchings in a tree, while the same coefficients in the distance matrix count certain obscurely labelled subforests [G&L] (Section I). However, as will be shown in this thesis, the characteristic polynomial of the distance matrix of a tree factors according to its automorphism group. In particular, the characteristic polynomial of the complete binary (and k-ary) tree has an explicit decomposition (Section IV).

The original questions I worked on involved generalizing some distance matrix theorems about trees to trees with weighted edges (Section II). It was hoped that weighting the edges would give some insight to the contribution of a single edge to the coefficients of the characteristic polynomial of the distance matrix of a tree. In fact, the contribution of a leaf can be isolated. Since a tree can be decomposed by removing one leaf at a time, it is possible, although time consuming, to build up the characteristic polynomial this way.

A lovely result about the determinant of the distance matrix of any graph is due to Graham, Hoffman and Hosoya. It states that the determinant can be computed using only the 2-connected pieces of the graph. In Section III, we attempt some generalizations of their theorem, most of which are disappointingly dull. However, it might be possible to prove Graham's unimodality conjecture (Conjectures V) for paths using this approach.

We conclude this introduction by mentioning what we believe to be the most interesting problem and conjecture still open. The problem is to try to factor the characteristic polynomials of distance matrices of other classes of trees than the binary trees. Since binary trees have so much symmetry, the automorphism group method of factoring the characteristic polynomial gives a lot of information. Other classes of trees probably will not respond with the same wealth of detail, but some common factors for common subtrees may be found. The most interesting open conjecture is whether binary trees are characterized by their distance matrix characteristic polynomials. We know that trees in general are not determined by their characteristic polynomials, but it seems unlikely that another tree could imitate the binary tree so well as to produce the same number of similar factors as the ones found in Section IV. If the binary trees are characterized by their characteristic polynomials, it would be interesting to know other classes of trees that are characterized by their polynomials.

## TABLE OF CONTENTS

## Section 0

For those who may not be completely familiar with graph theory terminology, we present a few definitions. Please be sure to read the last paragraph, since in it we define notation used throughout the text. For a more thorough discussion of basic graph concepts, we recommend [A]. A **directed graph G** is defined to be a (finite) set of vertices V(G) along with an edge set, E(G), which is a subset of V(G)xV(G). An **undirected graph** (usually just called graph) is an vertex set V(G) and an edge set E(G) which is a subset of all unordered pairs of vertices in V(G). The **degree** of a vertex is defined to be the number of edges it is contained in. If a vertex is in an edge e, it is said to be **adjacent** to e. If two edges share a common vertex v, they are said to be **incident** at v.

A **walk** of length r from vertex $v_1$ to vertex $v_{r+1}$ is a sequence of edges, $\{v_1,v_2\},\{v_2,v_3\}, \ldots, \{v_{r-1},v_r\}, \{v_r,v_{r+1}\}$. G is **connected** if for every pair of vertices u and v, there exists a walk from v to u. A **cycle** is a connected graph with every vertex of degree 2. A **forest** is a graph that contains no cycle and a **tree** is a forest that is connected. A **path** is a tree with every vertex of degree $\leq 2$. The **distance** between two vertices in a graph is the length of the shortest path between them.

A **leaf** is an edge in a tree that contains a vertex of degree 1. A **component** of a graph is one of its connected pieces. A graph is said to be **k-connected** if there exist k vertices whose removal disconnects the graph, but the removal of any k-1 vertices does not disconnect the graph. Note that a tree is always 1-connected. A **k-matching** is a graph with k edges, no pair of which are incident with no isolated vertices.

We define standard symbols used throughout the text. For any matrix M, let CP(M) be the characteristic polynomial of M. Label the vertices of a tree T with $v_1$, $v_2$, ..., $v_n$, and let $D(T) = (d_{ij})$ = the distance matrix of a tree, with $d_{ij}$ = the distance from vertex $v_i$ to vertex $v_j$. Let $\lambda^n + d_{n-2}\lambda^{n-2} + \ldots + d_0$ be the characteristic polynomial of D(T), CP(D(T)). Let $\Delta(T) = (\delta_{ij})$ be the weighted distance matrix, i.e. let the edges of a tree be labelled with $x_1$, $x_2$, ..., $x_{n-1}$ and $\delta_{ij}$ be the sum of variables from the $x_i$'s on the shortest path from vertex $v_i$ to vertex $v_j$. Let $CP(\Delta(T)) = \lambda^n + \delta_{n-2}\lambda^{n-2} + \ldots + \delta_0$. Let $A(T)=(a_{ij})$= the adjacency matrix of a tree, with $a_{ij}=1$ if vertices $v_i$ and $v_j$ are adjacent and 0 otherwise. Let $\lambda^n + a_{n-2}\lambda^{n-2} + \ldots + a_0$ be CP(A(T)).

## Section I

It is well known that the coefficients of CP(A(T)) satisfy: [CDS]

$$a_{n-k} = \begin{cases} (-1)^{n+k/2} L(k/2) & \text{if } k \text{ is even,} \\ 0 & \text{otherwise} \end{cases}$$

where L(i) = # of i-matchings in T. Graham and Lovász, in [G&L], investigate the coefficients of CP(D(T)), and show that the coefficients of the distance matrix satisfy a similar, although somewhat more complicated property.

Graham's and Lovász's result can be stated as:

**Theorem 1** Let T be a tree. Then $d_k(T) = (-1)^{n-1} 2^{n-k-2} \sum_F A_F^k N_F(T)$. where T has n vertices, F ranges over all forests with k-1,k,k+1 edges and no isolated vertices, and $A_F^k$ are integers depending only on F and k, and $N_F(T)$ is the number of forests F contained in T.

The coefficients $A_F^k$ can be written down explicitly in terms of the construction of the tree. My first result is a generalization of this. Let the edges of a tree be labelled with $x_1, x_2, \ldots, x_{n-1}$ so that $\delta_{ij}$ in the weighted distance matrix Δ(T) is now a sum of variables from the set of $x_i$'s. Let $CP(\Delta(T)) = \sum_{k=0}^{n} \delta_k(T)\lambda^k$.

**Theorem 2** $\delta_k(T) = (-1)^{n-1} 2^{n-k-2} \prod_i x_i \left\{ \sum_F \sum_{LF} \mathcal{A}_F^k \right\}$

where LF is a labelled copy of F in T and $\mathcal{A}_F^k$ is the weighted version of the A's. The $\mathcal{A}_F^k$'s and $A_F^k$'s will be defined at the end of the following proof.

This theorem reduces to Theorem 1 when all the $x_i$'s are equal to 1. The same explicit construction of the $\mathcal{A}_F^k$ can be given. We now proceed to give the proof of Theorems 1 and 2, by using a complicated labelling of the forests F contained in T. This proof is a straightforward generalization of the proof for the unweighted case in [G&L]. Much of the wording is taken directly from [G&L]; the whole proof is given for completeness.

**EXAMPLE** Let $P_2$ = o—o—o . Then $CP(A(P_2)) = -\lambda^3+2\lambda$, $CP(D(P_2)) = -\lambda^3+6\lambda+4$ and if the weighted version of $P_2$ = o$\underline{\ x_1\ }$o$\underline{\ x_2\ }$o, then $CP(\Delta(P_2)) = -\lambda^3 + 2\lambda(x_1^2 + x_1x_2 + x_2^2) + 2x_1x_2(x_1 + x_2)$

## PROOF OF THEOREMS 1 AND 2

The plan of the proof is fourfold:

1. The inverse $\Delta^{-1}$ of $\Delta=\Delta(T)$ is found.
2. We can use the relationship between the characteristic polynomial of a matrix and the characteristic polynomial of the inverse of the matrix to find $CP(\Delta(T))$ in terms of $CP(\Delta^{-1}(T))$.
3. The terms of the determinant expansion of $\Delta^{-1}-\lambda I$ are interpreted as counting the occurrences of certain labelled forests F in T.
4. The contribution each forest makes is determined.

Let us label the tree T as follows: Label the vertices $v_1,v_2,\ldots,v_n$ and let the first edge on the unique path from $v_i$ to $v_1$ be labelled with $x_{i-1}$ for $2\leq i\leq n-1$. Let us define the nxn matrix B by

$$B = \begin{pmatrix} 0 & x_1 & x_2 & \cdots & x_{n-1} \\ x_1 & -2x_1 & 0 & \cdots & 0 \\ x_2 & 0 & -2x_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \vdots & & 0 \\ x_{n-1} & 0 & 0 & \cdots\ 0 & -2x_{n-1} \end{pmatrix}$$

Let $N=(n_{ij})$ be the nxn matrix defined by $n_{ij} = 1$ if $d_{1j}= d_{1i}+ d_{ij}$

and 0 otherwise. Thus $n_{ij}=1$ iff the unique path from $v_1$ to $v_j$ contains $v_i$. Note that $n_{1i} = n_{ii} = 1$ for all i.

Lemma 1 $\Delta=N^T BN$

Proof of Lemma 1: Note that $BN = (c_{ij})$ where $c_{1j} = \delta_{1j}$ for $1\leq j\leq n$ and

$c_{ij}= x_{i-1}(1-2n_{ij})$ for $i\geq 2$. Multiplying by $N^T$, we get that

$$N^T BN = (c^*_{ij}) = \left\{ \sum_{k=1}^{n} n_{ki}c_{kj} \right\} = \left\{ n_{1j}\delta_{1j}+ \sum_{k=2}^{n} n_{ki}x_{k-1}(1-2n_{kj}) \right\} .$$

Now $\sum_{k=2}^{n} n_{ki}x_{k-1}(1-2n_{kj}) = \sum_{k=2}^{n} n_{ki}x_{k-1} - \sum_{k=2}^{n} 2n_{ki}n_{kj}x_{k-1}$. But

$$\sum_{k=2}^{n} n_{ki}x_{k-1} = \delta_{1i} \text{ and } \sum_{k=2}^{n} n_{ki}n_{kj}x_{k-1} = \begin{cases} 0 \text{ if } n_{ij}=n_{ji}=0 \\ \delta_{1i} \text{ if } n_{ij}=1 \\ \delta_{1j} \text{ if } n_{ji}=1 \end{cases} .$$

Hence $(c^*_{ij}) = (\delta_{ij})$.

Lemma 2 $N^{-1}= (\nu_{ij})$, where $\nu_{ij}= \begin{cases} 1 & \text{if } i=j \\ -1 & d_{ij}=1 \text{ and } d_{1j}=1+d_{1i} \\ 0 & \text{otherwise} \end{cases}$

Proof of Lemma 2: Note that $\nu_{ij} = -1$ iff $n_{ij} = 1$ and $v_i$ is adjacent to $v_j$. Since the $ij^{th}$ entry in the product $NN^{-1}$ is

$\sum_{k=1}^{n} n_{ik}\nu_{kj}$ , the only terms in the sum which have a nonzero contribution come from those k with both $n_{ik}\neq 0$ and $\nu_{kj}\neq 0$. If i=j then we must have k=i and the entry is 1. If $i\neq j$ then the only nonzero terms are for k=j with $n_{ij}\nu_{jj}= 1\cdot 1 = 1$ and $k=\kappa$, where $d_{\kappa j}=1$ and $d_{1j} = 1 + d_{1\kappa}$ with $n_{1\kappa}\nu_{\kappa j} = 1(-1) = -1$. Thus, for $i\neq j$ the entry is 0. Hence $NN^{-1}=I$.

<u>Lemma 3</u> $N^{-1}\left(N^{-1}\right)^T= \mathrm{diag}[d_1+1,d_2,d_3,\dots,d_n] - A(T)$ where $d_i$ = the degree of $v_i$.

Proof of <u>Lemma 3</u>: The $ij^{th}$ entry in the product, namely,

$\sum_{k=1}^{n}\nu_{ik}\nu_{kj}$, has the values $d_1+1$ if i=j=1, since all k with either k=1 of $v_k$ adjacent to $v_1$ contribute 1 to the sum; $d_i$ if i=j>1, since now we cannot take k=1; $-a_{ij}$ if $i\neq j$, for we cannot have $\nu_{ik}=\nu_{jk}\neq 0$ Thus, the only nonzero contribution can occur is $\nu_{ik}=-\nu_{jk}$, and $v_i$ is adjacent to $v_j$ so that $a_{ij}=1$. This proves the lemma.

<u>Lemma 4</u> Let $x_1+x_2+\dots+x_{n-1}=X$, and $X_i=X-x_i$. Then

$$B^{-1} = \frac{1}{2X}\begin{pmatrix} 4 & 2 & 2 & \dots & 2 \\ 2 & -X_1/x_1 & 1 & \dots & 1 \\ 2 & 1 & -X_2/x_2 & \dots & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 2 & 1 & 1 & \dots\ 1 & -X_{n-1}/x_{n-1} \end{pmatrix}$$

Proof of <u>Lemma 4</u> Multiply $B\cdot B^{-1}$.

<u>Lemma 5</u> $\Delta^{-1}=(\delta_{ij}^{-1})$ is given by

$$\delta_{ij}^{-1} = \frac{(2-d_i)(2-d_j)}{2X} + \begin{cases} -D_i/2 & \text{if } i=j \\ a_{ij}/x_{ij} & \text{if } i\neq j \end{cases}$$

where $D_i$ is the sum of the $1/x_j$ for the edges $x_j$ adjacent to $v_i$ and $\bar{x}_{ij}$ is the edge label of the edge adjacent to vertices $v_i$ and $v_j$. If $v_i$ and $v_j$ are not adjacent, then $x_{ij}$ is taken to be 1.

Proof of <u>Lemma 5</u> Use $\Delta^{-1} = N^{-1}B^{-1}\left(N^{-1}\right)^T$. The proof of the lemma is a straightforward calculation.

We now will use some elementary linear algebra to calculate the $CP\Delta^{-1}(T)$. Let $C(\lambda)$ be

$$\begin{pmatrix} -X & 2-d_1 & 2-d_2 & \cdots & 2-d_n \\ 2-d_1 & -2\lambda-D_1 & & a_{ij}/x_{ij} & \\ 2-d_2 & & \ddots & & \\ \vdots & & a_{ij}/x_{ij} & \ddots & \\ 2-d_n & & & & -2\lambda-D_n \end{pmatrix}$$

Now $C(\lambda)$ is $(n+1)x(n+1)$, so we label the first row and column of $C(\lambda)$ with 0 instead of 1, in order to use our previous notation.

Let $\det(C(\lambda)) = \sum_{k=0}^{n} c_k(T)\lambda^k$.

By performing elementary row and column transformations on $C(\lambda)$, (we multiply the $0^{th}$ row by $(2-d_i)/(n-1)$ and add the result to the $i^{th}$ row) we get that $\det[C(\lambda)] = -2^n X \cdot CP(\Delta^{-1})$. From elementary linear algebra, we know that $CP(\Delta)(\lambda) = (-\lambda)^n \det(\Delta)\cdot CP(\Delta^{-1})(\lambda^{-1})$. We also need

<u>Lemma 6</u> [GHH] $\det(\Delta(T)) = (-1)^{n-1}(2)^{n-2}\cdot X \cdot \prod_{i=1}^{n-1} x_i$.

Proof of <u>Lemma 6</u> The proof appears in Section III.

Using Lemma 6, we have $CP(\Delta)(\lambda) = -2^{n-2}X\cdot\lambda^{n}\prod_{i=1}^{n-1}x_i\cdot CP(\Delta^{-1})(\lambda^{-1})$.

Hence $\delta_k(T) = (1/4)\prod_{i=1}^{n-1}x_i\cdot c_{n-k}(T)$.

If we expand $\det[C(\lambda)]$ and collect the terms which contribute to the $\lambda^{n-k}$th term, we find

$$c_{n-k}(T) = (-2)^{n-k}\sum_{i_1,i_2,\dots,i_k}\det\begin{pmatrix} -X & 2-d_{i_1} & \cdots & 2-d_{i_k} \\ 2-d_{i_1} & -D_{i_1} & & a_{i_s i_r}/x_{i_s i_r} \\ \vdots & & \ddots & \\ 2-d_{i_k} & a_{i_s i_r}/x_{i_s i_r} & & -D_{i_k} \end{pmatrix}$$

where the sum ranges over all choices of $1\leq i_1<\dots<i_k\leq n$.

Let us examine the expansion of the general determinant

$$\det\begin{pmatrix} -X & 2-d_{i_1} & \cdots & 2-d_{i_k} \\ 2-d_{i_1} & -D_{i_1} & & a_{i_r i_s}/x_{i_r i_s} \\ \vdots & & \ddots & \\ 2-d_{i_k} & a_{i_r i_s}/x_{i_r i_s} & & -D_{i_k} \end{pmatrix} \qquad (1)$$

where we label the rows and columns with $\{0,1,\dots,k\}$. An important observation is that the only permutation choices from the above matrix which can contribute nonzero terms to the determinant correspond to permutations

$\pi=(0 j_1\ j_2\cdots j_s)(j_{s+1}\ j_{s+2})(j_{s+3}\ j_{s+4})\cdots(j_{s+2m-1}\ j_{s+2m})(j_{s+2m+1})\cdots(j_k)$

This follows at once from the fact that since T contains no cycles, the only nontrivial cycles $\pi$ can have either involve row 0 and column 0 or have length 2. Furthermore, all the terms $a_{j_1 j_2}, a_{j_2 j_3},\dots,a_{j_{s-1} j_s}$ must be 1.

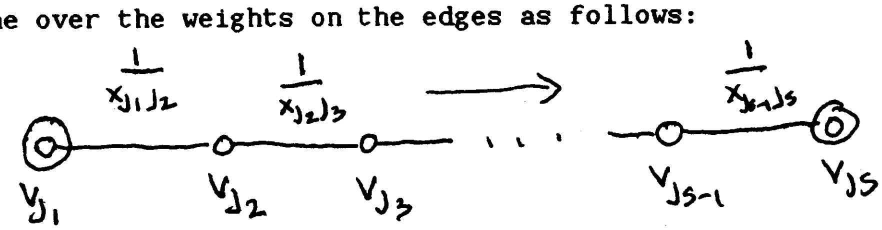
Let $\alpha_{ij}=\frac{a_{ij}}{x_{ij}}$. The permutation $\pi$ above corresponds to the term (2)

$$(2-d_{j_1})\alpha_{j_1j_2}\cdots\alpha_{j_{s-1}j_s}(2-d_{j_s})(\alpha_{j_{s+1}j_{s+2}}\alpha_{j_{s+2}j_{s+1}})\cdots(-D_{j_{s+2m+1}})\cdots(-D_{j_k})$$

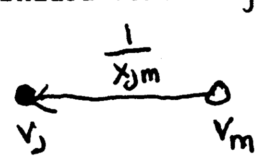
in the expansion of (1). When s=0, this has the slightly different form $-X\cdot(\alpha_{j_{s+1}j_{s+2}}\alpha_{j_{s+2}j_{s+1}})\cdots(-D_{j_{s+2m+1}})\cdots(-D_{j_k})$. We may expand the term in (2) into three similar terms formed by replacing $(2-d_{j_1})(2-d_{j_s})$ by $d_{j_1}d_{j_s}-2(d_{j_1}+d_{j_s})+4$.

Next, we interpret the individual terms in the expansion of the determinants in the above $c_{n-k}$ as enumerating certain subforests F of T in which the vertices and edges of F have been marked in various ways.

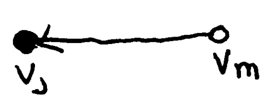
(i) For the factor $\alpha_{j_1j_2},\alpha_{j_2j_3},\ldots,\alpha_{j_{s-1}j_s}$ we distinguish the endpoints and the direction on the path in T from $v_{j_1}$ to $v_{j_s}$ (if s>2) and one over the weights on the edges as follows:



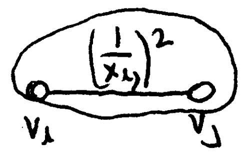
(ii)(a) For the factor $D_j$, we mark an edge of T incident to $v_j$ with an arrow pointing to the shaded vertex $v_j$ and one over the weight of the edge chosen:



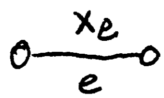
(ii)(b) For the factor $d_j$, we mark an edge of T incident to $v_j$ with an arrow pointing to the shaded vertex $v_j$:



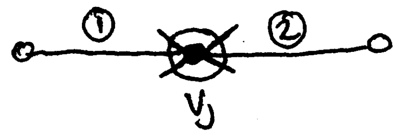
(iii) For the factor $(\alpha_{ij}\alpha_{ji})$ we distinguish the edge $\{v_i, v_j\}$ in T and mark it with one over its weight squared:




(iv) For the factor X, we mark an edge with the symbol e and its weight:

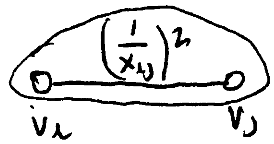



(v) For the factor $d_j d_j$ (which will occur only when s=1), we mark one edge incident to $v_j$ with an arrow to $v_j$ and a symbol 1 and we mark one edge (possibly the same edge) with an arrow to $v_j$ and a symbol 2; also, we circle and shade $v_j$.

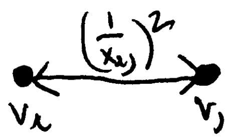



The terms in (2) now correspond exactly to the number of ways T can be marked according to rules just given. Of course, one must keep in mind the fact that degeneracies may occur; e.g., some edges of T may receive several marks. The value of (1) is now given by enumerating all possible ways of marking T according to (i) through (v) and summing the appropriate signed expressions over all choices of $1 \leq j_1 < \ldots < j_k \leq n$.

Of course, the terms of determinant (1) also have signs attached to them. Specifically each term with the cycle structure of $\pi$ defined above has an additional sign factor of $(-1)^{s+m}$.

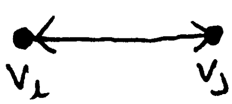
A considerable simplification now results from the following observation. For each marking of T which contains an edge marked by (iii), i.e.: 

(because of a factor $(\alpha_{ij}\alpha_{ji})$) there is another marking of T which is identical except for the edge $\{v_i,v_j\}$, now (degenerately) marked by

(ii)(a) as 

Furthermore, the corresponding terms in the expansion of **(1)** from which the two markings come have opposite signs. This is obvious, since the two permutations differ only in that the factor $(\alpha_{ij}\alpha_{ji})$ in one is replaced by $(-D_i)(-D_j)$ in the other and such a change certainly changes the sign of the permutation. Hence, all the contributions from the markings of the type in (iii) are canceled out by all the markings in which the edge $\{v_i,v_j\}$ has two arrows, one to $v_i$ and one to $v_j$, and for which $D_i$ and $D_j$ have been selected from the diagonal of **(1)**.

Thus, we may henceforth restrict our consideration to permutations $\pi$ for which m=0: $\pi=(0j_1\cdots j_s)(j_{s+1})\cdots(j_k)$ **(3)**, provided the edges $\{v_i,v_j\}$ in T marked as



come only from the factors $(2-d_i)(2-d_j)$ of a term i.e., have no weight label. However, since for any permutation $\pi$ above there is at most one cycle $(0j_1\cdots j_s)$ containing 0, in the corresponding markings of T, at most one edge can have arrows at each of its endpoints.

The specific terms of the determinant which come from the shortened $\pi$ above are

I. $s=0$: $-X\cdot(-D_{j_1})\cdots(-D_{j_k})$, sign $\pi=1$

II. $s=1$: $(2-d_{j_1})(2-d_{j_1})(-D_{j_2})\cdots(-D_{j_k})$, sign $\pi=-1$

We split this into the sum of the three terms:

(i) $d_{j_1}d_{j_1}(-D_{j_2})\cdots(-D_{j_k})$

(ii) $-4d_{j_1}(-D_{j_2})\cdots(-D_{j_k})$

(iii) $4(-D_{j_2})\cdots(-D_{j_k})$

III. $s\geq 2$: $(2-d_{j_1})\alpha_{j_1j_2}\cdots\alpha_{j_{s-1}j_s}(2-d_{j_s})(-D_{j_{s+1}})\cdots(-D_{j_k})$, sign $\pi=(-1)^s$

We also split this into the sum of three terms:

(i) $(d_{j_1}-1)\alpha_{j_1j_2}\cdots\alpha_{j_{s-1}j_s}(d_{j_s}-1)(-D_{j_{s+1}})\cdots(-D_{j_k})$

(ii) $-((d_{j_1}-1)+(d_{j_s}-1))\alpha_{j_1j_2}\cdots\alpha_{j_{s-1}j_s}(-D_{j_{s+1}})\cdots(-D_{j_k})$

(iii) $\alpha_{j_1j_2}\cdots\alpha_{j_{s-1}j_s}(-D_{j_{s+1}})\cdots(-D_{j_k})$

Our next task is to examine the number of ways a given subforest F of T can be marked so as to contribute to the nonzero terms in I and II. If F is a forest with connected components $C_1,\ldots,C_t$ (which of course are trees), we define

$|F|$=the number of vertices of F,

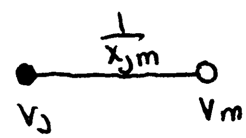
$||F||$=the number of edges of F,

$\rho(F)=\prod_{i=1}^{t}|C_i|$, and $x_F=\sum_{x_i\subset F} x_i$

We let $\Gamma_k$ denote the set of all forests having no isolated vertices and exactly k edges. For the empty forest $F^*$, we set $\pi(F^*)=1$.

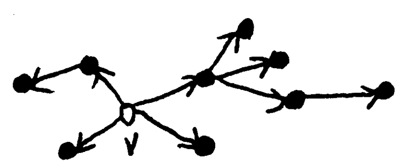
Since each factor $d_{j_i}$, $D_{j_i}$ and $\alpha_{ij}$ corresponds to the marking of a unique edge of T, with the exception of $d_{j_1}$ and $d_{j_s}$ in II which may degenerate, it follows that we must have $||F||=k+1,k,$or $k-1$.

Let F be an arbitrary fixed subforest of T with components $C_1,\ldots,C_t$ and no isolated points. We wish to determine in how many ways T may be marked according to the conditions given above so that the marked edges are exactly the edges of F. Because of the restrictions on marking T, it follows that all $C_i$ except possibly one, which we denote by $\bar{C}^*$, have all edges marked according to (ii)(a), i.e., as




We say that $C_i$ is marked normally in this case. The number of ways $C_i$ can be marked normally is just $|C_i| \cdot \frac{1}{x_{C_i}}$, the number of vertices of $C_i$ times one over the product of the edge weights of $C_i$. This is because each vertex of such a $C_i$, except for exactly one vertex v, must have exactly one edge with an incoming arrow. Now v has all edges with outgoing arrows. Thus, v serves as a "source" and the direction of all other arrows are determined. (See Figure 1.) Hence, it suffices to determine the number of ways the exceptional component $C^*$ can be marked. Each edge in Figure 1 is labelled with one over its variable label, $\frac{1}{x_{edge}}$, the variable labels being taken from our original labelling of the edges of T. These labels are not present in Figure 1 to avoid confusing the eye.

Figure 1



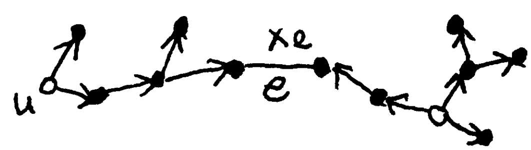

As we have noted, F can have only k+1,k, or k-1 edges. We treat the three cases separately. We first interpret the absolute value of the terms and then determine the appropriate signs. In the first case, with $||F|| = k+1$, there can be no multiply labelled edges. There are k+1 labels and k+1 edges and each edge must receive a label.

$||F||=k+1$

I. $X\cdot(D_{j_1})(D_{j_2})\cdots(D_{j_k})$ See Figure 2. If the edge e corresponding to the choice for the factor X is erased from $C^*$, the two resulting components can be arbitrarily marked normally, that is, using (ii). Thus, for each choice of "source" vertices u and v in $C^*$, there are d(u,v)=the unweighted distance between u and v in the tree T possible locations of the edge e. Therefore, there are exactly

$$\sum_{\{u,v\}\subseteq C^*} d(u,v)$$

ways of marking $C^*$ in this case. The weight of $C^*$ is always $x_e^2/x_{C^*}$. This is because all the edges except e are weighted with one over their usual label, $\frac{1}{x_{edge}}$, and e is weighted with $x_e$.

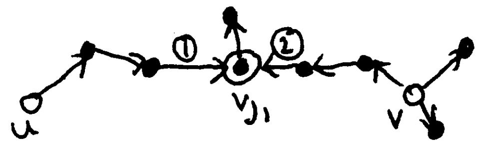
Figure 2 


Let $\delta(u,v)$ = the set of the edge weights on the path from u to v. Since the sign of the permutation $\pi$ in (3) is +1, the total contribution to the determinant is

$$(-1)^{k+1}\,\frac{\rho(F)}{x_F}\;\sum_{i=1}^{t}\frac{1}{|C_i|}\;\sum_{\{u,v\}\subseteq C_i}\;\sum_{x_r\in\delta(u,v)} x_r^2$$

II(i). $d_{j_1}d_{j_1}(-D_{j_2})\cdots(-D_{j_k})$. See Figure 3. Since $||F||=k+1$, there are no multiply marked edges. We use an argument similar to that in the preceding case (where an extra factor of 2 comes from the labelling of the edges with 1 or 2). All the edges except those labelled 1 and 2 are labelled with one over their usual variable.

Figure 3

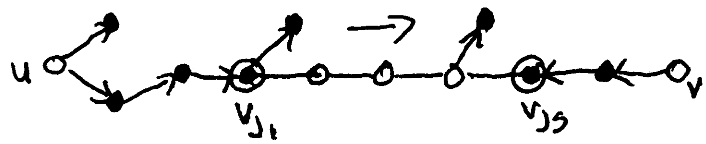



We obtain a total contribution in this case of

$$(-1)^{k}2\ \frac{\rho(F)}{x_F}\sum_{i=1}^{t}\frac{1}{|C_i|}\sum_{\{u,v\}\subseteq C_i}\ \sum_{\{x_r,x_s\}\subset\delta(u,v),\ x_r\neq x_s \text{ are incident in } T} x_r x_s$$

III(i). $(d_{j_1}-1)\alpha_{j_1 j_2}\cdot\cdot\alpha_{j_{s-1} j_s}(d_{j_s}-1)(-D_{j_{s+1}})\cdots(-D_{j_k})$. Now on the path between the sources x and y we must choose the two points $v_{j_1}$ and $v_{j_s}$ as well as a direction. All the edges except the neighbor of $v_{j_1}$ on the unique path from u to $v_{j_1}$, and the neighbor of $v_{j_s}$ on the unique path from v to $v_{j_s}$ have the usual weight of $\frac{1}{x_{edge}}$.

Figure 4




Thus, the total contribution in this case is

$$(-1)^{k}2\frac{\rho(F)}{x_F}\sum_{i=1}^{t}\frac{1}{|C_i|}\sum_{\{u,v\}\subseteq C_i}\ \sum_{x_r\neq x_s \text{ not incident in } T;\ \{x_r,x_s\}\subset\delta(u,v)} x_r x_s$$

The remaining cases II(i), (iii), and III(ii), (iii) cannot contribute for $||F||=k+1$. Combining the above results, we get

<u>Lemma 7</u>: If $||F||=k+1$, then

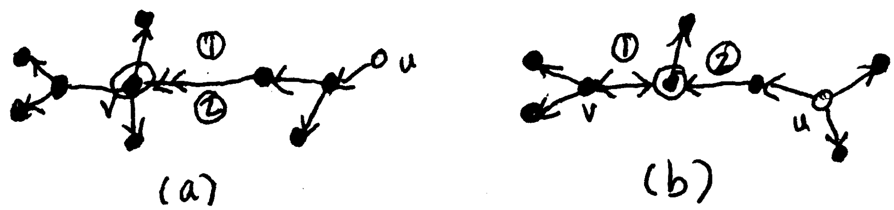
$$\Delta_F^k = \frac{\rho(F)}{x_F}\sum_{i=1}^{t}\frac{1}{|C_i|}\sum_{\{u,v\}\subseteq C_i}\left\{\sum_{x_r\subset\delta(u,v)} x_r^2 - \sum_{\{x_r,x_s\}\subset\delta(u,v)} x_r x_s\right\}$$

**$||F||=k$**

I. $-X\cdot(-D_{j_1})\cdots(-D_{j_k})$. The multiply marked edge must be an edge which has both an arrow and the symbol e. All components are initially marked normally. Then an arbitrary edge of F is selected for e. This edge contributes $\frac{1}{x_e}\cdot x_e = 1$ as its weight. The total contribution is therefore

$$(-1)^{k+1}\,\frac{\rho(F)}{x_F}\sum_{x_r\subset F} x_r$$

II(i) $d_{j_1}d_{j_1}(-D_{j_2})\cdots(-D_{j_k})$. There are two ways an edge can be lost. They are shown in Figure 5. In both cases, we must sum over (u,v) and (v,u) since the sums are not symmetric in u and v.

Figure 5 


In 5(a), each edge contributes $\frac{1}{x_{edge}}$ except the one labelled with 1 and 2. The total contribution is

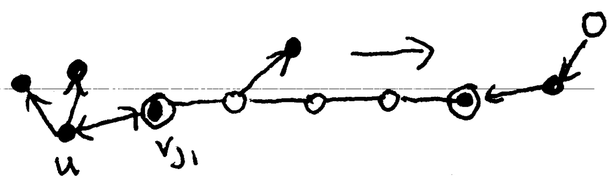
$$(-1)^k\,\frac{\rho(F)}{x_F}\sum_{i=1}^{t}\frac{1}{|C_i|}\sum_{(u,v)\subseteq C_i;\, x_{uv}} x_{uv}$$ is the edge on the path from u to v incident to v.

$$= \quad (-1)^k\,\frac{\rho(F)}{x_F}\sum_{x_r\subset F} x_r$$

In figure 5(b), the points u and v cannot be adjacent. Also, we have a factor of 2 corresponding to the assignment of 1 and 2. Each edge contributes $\frac{1}{x_{edge}}$ except the edge labelled 2. Thus, the total contribution in this case is

$$(-1)^k \frac{2\rho(F)}{x_F} \sum_{i=1}^{t} \frac{1}{|C_i|} \sum_{(u,v)\subseteq C_i;\ d(u,v)>=2;\ x_{uv} \text{ is the second edge on the path from v to u.}} x_{uv}$$

III(i). $(d_{j_1}-1)\alpha_{j_1 j_2}\cdot\cdot\alpha_{j_{s-1} j_s}(d_{j_s}-1)(-D_{j_{s+1}})\cdots(-D_{j_k})$. The factor $(d_{j_1}-1)$ is interpreted as choosing any edge with no weight adjacent to $v_{j_1}$ except the one on the path between $v_{j_1}$ and $v_{j_s}$ (with $(d_{j_s}-1)$ interpreted similarly). The only possibility for marking C* is shown in Figure 6. Hence the only edge without a weight is the edge corresponding to $(d_{j_s}-1)$. We must have $d(u,v)\geq 3$, since $s\geq 2$. Once u and v and the direction are chosen, there are d(u,v) - 2 choices for $v_{j_s}$. Each edge contributes $\frac{1}{x_{edge}}$ except the neighbor of $v_{j_s}$ on the unique path to v.

Figure 6 

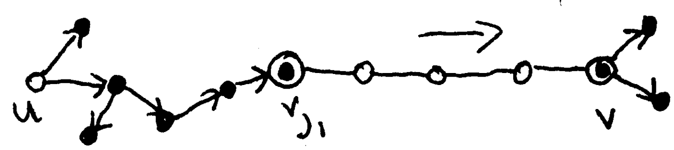

Thus, in this case the contribution is

$$(-1)^k \frac{2\rho(F)}{x_F} \sum_{i=1}^{t} \frac{1}{|C_i|} \sum_{\substack{(u,v)\subseteq C_i;\\ d(u,v)>=3}} \sum_{x_r \subset \delta(u,v);\ x_r \text{ is not the 1st or 2nd edge on the path from u to v}} x_r$$

II(ii). $4d_{j_1}(-D_{j_2})\cdots(-D_{j_k})$. This "term" has the property that it appears k times in the expansion of the determinant, once for each choice of the small term $d_{j_1}$. Since there are just k factors in it and $||F||=k$ in this case, no edges are lost, all components are marked normally but the term $d_{j_1}$ contributes no weight to the total. The total contribution is

$$(-1)^{k+1} \frac{4\rho(F)}{x_F} \sum_{x_r \subset F} x_r$$

III(ii). $-((d_{j_1}-1)+(d_{j_s}-1))\alpha_{j_1 j_2}\cdots\alpha_{j_{s-1} j_s}(-D_{j_{s+1}})\cdots(-D_{j_k})$ The markings corresponding to this case are shown in Figure 7. As before, we must have $\bar{d}(u,v)\geq 2$. Each edge contributes the usual except the neighbor of $v_{j_1}$ on the path from u to $v_{j_1}$.

Figure 7 

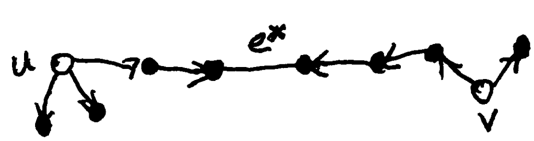

The contribution is readily calculated to be

$$(-1)^{k+1}\frac{2\rho(F)}{x_F}\sum_{i=1}^{t}\frac{1}{|C_i|}\sum_{(u,v)\subseteq C_i}\sum_{x_r\in\delta(u,v);\ x_r \text{ is not incident to } v} x_r$$

II(iii). $4(D_{j_2})\cdots(D_{j_k})$ At first sight it would appear that there are no contributions to $||F||=k$ from this case. However, it must be recognized that this term actually occurs n-k+1 times in the expansion of $\delta^{-1}_{n-k}(T)$. We can write this as

$$4(n-k+1)(D_{j_2})\cdots(D_{j_k}) = 4(n-k)(D_{j_2})\cdots(D_{j_k}) + 4(D_{j_2})\cdots(D_{j_k})$$

We interpret the term $4(n-k)(D_{j_2})\cdots(D_{j_k})$ as selecting distinct edges incident to $v_{j_2},\ldots,v_{j_k}$ (as usual) together with another distinct edge e* with no weight (since T has n-1 edges altogether). The corresponding marking is shown in Figure 8. Each edge has weight $\frac{1}{x_{edge}}$ except e*.

Figure 8 


This therefore contributes

$$(-1)^k \frac{4\rho(F)}{x_F} \sum_{i=1}^{t} \frac{1}{|C_i|} \sum_{\{u,v\}\subseteq C_i} \sum_{x_r \subset \delta(u,v)} x_r$$

There are no other contributions to $||F||=k$. We may now sum all the preceding expressions for the case $||F||=k$ to obtain the following result.

Lemma 8 If $||F||=k$, then

$$\Delta_F^k = \frac{4\rho(F)}{x_F}\left\{ \sum_{x_r \subset F} x_r - \sum_{i=1}^{t} \frac{1}{|C_i|} \sum_{\{u,v\}\subseteq C_i} \sum_{x_r \subset \delta(u,v)} x_r \right\}$$

**$||F||=k-1$**

There are no contributions here from I. In this case, every edge gets a label, so we don't have to cancel out any $x_i$'s, since only unlabelled edges can overlap.

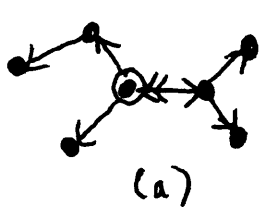
II(i). $d_{j_1} d_{j_1} (D_{j_2}) \cdots (D_{j_k})$. There are two possibilities here. They are shown in Figure 9. In Figure 9(a), an edge of C* is chosen and one end is distinguished. Thus, this case contributes

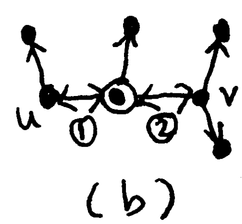
$$(-1)^k \frac{2\rho(F)}{x_F} \sum_{i=1}^{t} \frac{||C_i||}{|C_i|}$$

In Figure 9(b), a pair of points u,v with d(u,v)=2 is chosen and the two edges between them are ordered. Hence, this case contributes

$$(-1)^k \frac{2\rho(F)}{x_F} \sum_{i=1}^{t} \frac{1}{|C_i|} \sum_{\{u,v\}\subseteq C_i;\, d(u,v)=2} 1$$

Figure 9

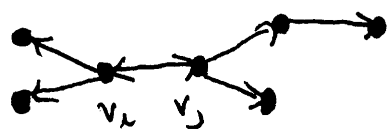


II(ii). $4d_{j_1}(D_{j_2})\cdots(D_{j_k})$. The marking shown in Figure 10 can contribute to only two terms, namely, $4(D_{j_1})\cdots d_i\cdots(D_j)\cdots(D_{j_k})$ and $4(D_{j_1})\cdots(D_i)\cdots d_j\cdots(D_{j_k})$, since the parenthesized $D_i$'s cannot place arrows on the same edge. Hence, to mark C*, we simply choose a distinguished edge. The factor of 2 comes from the two terms to which this marking contributes. The resulting expression is therefore

$$(-1)^{k+1}\frac{8\rho(F)}{x_F}\sum_{i=1}^{t}\frac{||C_i||}{|C_i|}$$

Figure 10

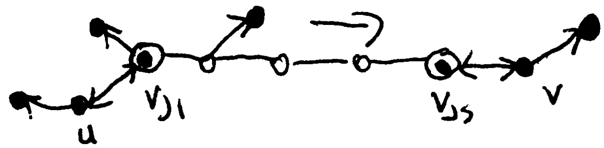



II(iii). $4(D_{j_2})\cdots(D_{j_k})$. From the discussion of case II(iii) for $||F||=k$, we may use the expansion discussed there. In particular, the second term $4(D_{j_2})\cdots(D_{j_k})$, now summed just over $1\leq j_2<\cdots<j_k\leq n$ results in a contribution of

$$(-1)^{k}\frac{4\rho(F)}{x_F}$$

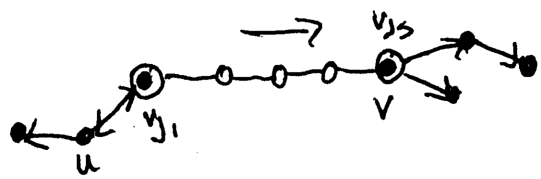
III(i). $(d_{j_1}-1)\alpha_{j_1j_2}\cdots\alpha_{j_{s-1}j_s}(d_{j_s}-1)(-D_{j_{s+1}})\cdots(-D_{j_k})$. The markings of C* contributing to $||F||=k-1$ are shown in Figure 11. Since $s\geq 2$, we must have $d(u,v)\geq 3$. The contribution here is

$$(-1)^{k}\frac{2\rho(F)}{x_F}\sum_{i=1}^{t}\frac{1}{|C_i|}\sum_{\{u,v\}\subseteq C_i;\,d(u,v)>=3}1$$

Figure 11 


III(ii). $((d_{j_1}-1)+(d_{j_s}-1))\alpha_{j_1j_2}\cdots\alpha_{j_{s-1}j_s}(-D_{j_{s+1}})\cdots(-D_{j_k})$. The corresponding marking is shown in Figure 12. We must have $d(u,v)\geq 2$. Thus, we obtain a contribution of

$$(-1)^{k+1}\frac{4\rho(F)}{x_F}\sum_{i=1}^{t}\frac{1}{|C_i|}\sum_{\{u,v\}\subseteq C_i;d(u,v)>=2}1$$

Figure 12 


III(iii). $\alpha_{j_1j_2}\cdots\alpha_{j_{s-1}j_s}(-D_{j_{s+1}})\cdots(-D_{j_k})$. We show the marking in Figure 13. The contribution is

$$(-1)^{k}\frac{2\rho(F)}{x_F}\sum_{i=1}^{t}\frac{1}{|C_i|}\sum_{\{u,v\}\subseteq C_i}1 \quad = \quad (-1)^{k}\frac{\rho(F)}{x_F}\,||F||$$

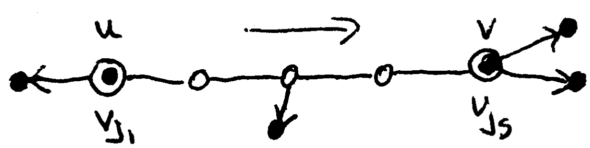


Figure 13

By combining all the expressions we obtain

<u>Lemma 9</u> If $||F||=k-1$, then

$$\delta_F^k = \frac{4\rho(F)}{x_F}\left\{\sum_{i=1}^{t}\frac{||C_i||}{|C_i|} - 1\right\}$$

We combine lemmas 6,7and 8 to get

**THEOREM 2** Let T be a tree with $n\geq 2$ vertices and distance matrix $\Delta(T)$. Let $\Gamma_k$ be the family of forests of k edges. Let LF be a labelled copy of F in T. If we write $CP(\Delta(T)) = \sum_{k=0}^{n}\delta_k(T)\lambda^k$, then

$$\delta_k = (-1)^{n-1} 2^{n-k-2} \prod_{r=1}^{n-1} x_r \Bigg\{$$

$$\sum_{F\subset\Gamma_{k+1}} \sum_{LF\subset T} \frac{\rho(F)}{x_F} \sum_{i=1}^{t} \frac{1}{|C_i|} \sum_{\{u,v\}\subseteq C_i} \left\{ \sum_{x_r\subset\delta(u,v)} x_r^2 - \sum_{\{x_r,x_s\}\subset\delta(u,v)} x_r x_s \right\}$$

$$\sum_{F\subset\Gamma_k} \sum_{LF\subset T} \frac{4\rho(F)}{x_F} \left\{ \sum_{x_r\subset F} x_r - \sum_{i=1}^{t} \frac{1}{|C_i|} \sum_{\{u,v\}\subseteq C_i} \sum_{x_r\subset\delta(u,v)} x_r \right\}$$

$$\sum_{F\subset\Gamma_{k-1}} \sum_{LF\subset T} \frac{4\rho(F)}{x_F} \left\{ \sum_{i=1}^{t} \frac{||C_i||}{|C_i|} - 1 \right\} \Bigg\}$$

**Example** $d_0(T) = (-1)^{n-1} 2^{n-2} (n-1)$ for all trees T

$$d_1 = (-1)^{n-1} 2^{n-3} (4N_1 + 2N_2 + 4N_3 - 4)$$

where $N_1$ is the number of pairs of disjoint edges, $N_2$ is the number of paths on 2 edges and $N_3$ is the number of edges.

**Example** $\delta_0 = (-1)^{n-1} 2^{n-2} \prod x_i \left\{ \sum x_i \right\}$

$$\delta_1 = (-1)^{n-1} 2^{n-3} \prod_i x_i \left\{ 2\Big(\sum_i x_i\Big) \Big(\sum_i \frac{1}{x_i}\Big) - \sum_{i=1}^{n} (d_i - 2)^2 \right\}$$

**Section II**

If we could isolate the contribution of a leaf of T to CP(Δ(T)), then we could inductively build up CP(Δ(T)) by peeling off one leaf of T at a time. Thus, one would hope that the CP(Δ(T)) would contain the CP(Δ(T*)), whenever T* is T minus a leaf. However, the coefficients of the bigger tree contain more information than those of the smaller tree and include forests in which the leaf is both an isolated and not an isolated edge.

For example, let $P_3 = o \overset{x_1}{\text{---}} o \overset{x_2}{\text{---}} o \overset{x_3}{\text{---}} o$. Then CP(Δ($P_3$))=

$$\lambda^4 - \lambda^2(3x_1^2 + 4x_2^2 + 3x_3^2 + 4x_1x_2 + 4x_2x_3 + 2x_1x_3)$$
$$- 2\lambda(2(x_1^2x_2 + x_1^2x_3 + x_2^2x_1 + x_2^2x_3 + x_3^2x_1 + x_3^2x_2) + 4x_1x_2x_3)$$
$$-4x_1x_2x_3(x_1 + x_2 + x_3).$$

Now if we take the partial derivative of the coefficient of λ with respect to the leaf $x_3$ and set $x_3$=0, we get

$\frac{-1}{2}\frac{\delta}{\delta x_3}\left( \delta_1(P_3)\right)\Big|_{x_1=0} = 2(x_1^2 + x_2^2 + 2x_1x_2)$, but $\delta_1(P_2) = 2(x_1^2 + x_1x_2 + x_2^2)$.

However, if we set $x_1$=0, for $x_1$ a leaf, we can say something inductive about CP(Δ(T)) if we know CP(Δ(G)) for graphs which are no longer trees. For any matrix M, let M(i,j) be M with the $i^{th}$ row and $j^{th}$ column deleted. If M=Δ(T), then Δ(T)(i,i) is the distance matrix of the graph $G_i$ with its $i^{th}$ vertex $v_i$ replaced by a complete graph $K_i$ on the neighbors of $v_i$, and the edge from $v_j$ to $v_k$ in the complete graph labelled with the sum $x_{ij} + x_{ik}$.

**Theorem 3** Let $x_1$ be a leaf and $v_1$ its vertex of degree 1, $v_2$ its other vertex. Let $T_1$ be the subtree of T without the edge $x_1$. Let G be the graph whose weighted distance matrix is given by $\big(\Delta(T)(1,1)\big)(2,2)$.

Then $CP(\Delta(T))\big|_{x_1=0} = -2\lambda CP(\Delta(T_1)) - \lambda^2 CP(\Delta(G_2))$

**Proof of Theorem 3** Once we have set $x_1$ to zero, the first and second rows are the same except in the first two entries:

$$\begin{matrix} -\lambda & 0 & \delta_{23} & \delta_{24} \cdots \delta_{2n} \\ 0 & -\lambda & \delta_{23} & \delta_{24} \cdots \delta_{2n} \end{matrix}$$

. (Since the matrix is symmetric, the first and second columns are also the same, except in the first two entries.) We subtract the second row from the first row and the second column from the first column to obtain in the first two rows:

$$\begin{matrix} -2\lambda & \lambda & 0 & 0 & \cdots & 0 \\ \lambda & -\lambda & \delta_{23} & \delta_{24} & & \delta_{2n} \end{matrix}$$

Expanding by the first row gives us $-2\lambda CP(\Delta(T_1)) - \lambda^2 \det(M_3)$, where $M_3$ is the n-2 x n-2 submatrix of D(T) indexed by rows 3 to n and columns 3 to n. It is easily seen that $M_3$ is the weighted distance matrix of $G_2$ above.

Theorem 3 above gives the most information when, not only is $x_1$ a leaf, but its vertex $v_2$ has degree only 2. In this case, $G_2$ is a tree and each coefficient of $CP(\Delta(T))$ "contains" the coefficient of of $CP(\Delta(T-x_1))$ and of $CP(\Delta(T-\{v_1,v_2\}))$.

**Example** When T is a path, we get the most flavor of the theorem.

$$CP(\Delta(o \overset{x_1}{\_\_\_} o \overset{x_2}{\_\_\_} o \overset{x_3}{\_\_\_} o \;\cdots\; o \overset{x_k}{\_\_\_} o))\Big|_{x_1=0} =$$

$$-2\lambda CP(\Delta(o \overset{x_2}{\_\_\_} o \overset{x_3}{\_\_\_} o \cdots o \overset{x_k}{\_\_\_} o) - \lambda^2 CP\Delta(o \overset{x_3}{\_\_\_} o \cdot\cdot o \overset{x_k}{\_\_\_} o))$$

We can describe the contribution to $CP(\Delta(T))$ of a single leaf $x_1$ by writing $CP(\Delta(T))$ as a polynomial in $x_1$, but we lose our inductive interpretation of the terms. In the following theorem, we substitute determinants which are no longer the determinants of subtrees of T, but instead can only be expressed in terms of the minors of D(T). However, it is possible to compare two characteristic polynomials using this theorem to say exactly if they are equal.

Let $x_1$ be a leaf and $v_1$ its vertex of degree 1. Let $S = \left(\Delta(T) - \lambda I\right)(1,1)$. So S has no entries containing $x_1$. For the sake of clarity, we label the rows and columns of S from the set {2,..,n}, instead of the usual. Then $CP(\Delta(T)) = \det(S)\left(-\lambda - z^T S^{-1} z\right)$ where $z^T = \left[\delta_{12}\ \delta_{13} \ldots \delta_{1n}\right]$. This is clear, since it is just expanding the determinant by the first row. Let $D_2 = \left[\delta_{22}\ \delta_{23} \ldots \delta_{2n}\right]$ and $J_1 = \left[x_1\ x_1 \ldots x_1\right]$. So $z = J_1 + D_2$. Let $S^{-1} = (s^{-1}_{ij})$. Let $cof(S) = \det(S)\sum_{i,j} s^{-1}_{ij}$. Then we have

**Theorem 4** $CP(\Delta(T)) = -\lambda\det(S) - \det(S)D_2^T S^{-1} D_2$

$$-2x_1\left(\det(S) + \lambda\sum_{j=2}^{n}(-1)^{2+j}\det\left(S(2,j)\right)\right) - x_1^2 cof(S).$$

$$= -2\lambda\det(S) - \lambda^2\det(S(2,2))$$

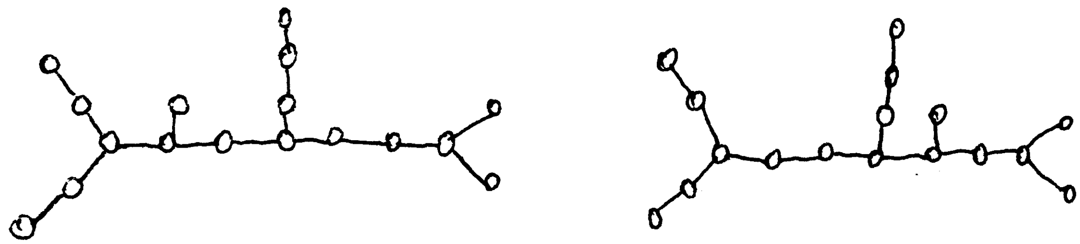
$$-2x_1\left(\det(S) + \lambda\sum_{j=2}^{n}(-1)^{2+j}\det\left(S(2,j)\right)\right) - x_1^2 cof(S).$$

From this we see exactly the contribution of $x_1$.

**Proof of Theorem 4** The proof is simply expanding the term $z^T S^{-1} z$ and realizing that $D_2^T S^{-1} D_2 = \lambda + \lambda^2 \frac{\det(S(2,2))}{\det(S)}$.

**Cospectral Trees**

Two trees are said to be cospectral if they have the same CP(D(T)). The smallest pair of cospectral trees are the following: (due to McKay [McK]).



The removal of a leaf from each of these gives the same tree. McKay has shown that there are infinitely many cospectral pairs as the number of vertices goes to infinity. All the pairs in his paper contain copies of the trees above. The following theorem gives necessary and sufficient conditions for two trees that differ only by a leaf to be cospectral. The conditions are a direct result of Theorem 4.

**Theorem 5** Let T be a tree with n vertices as usual and $v_1, v_2$ be vertices of T. Let $T_i$ be T with a leaf added at $v_i$ and the leaf vertex of degree 1 labelled $x_{n+1}$, since $T_i$ has n+1 vertices. Let $D(T) = (d_{ij})$ be the distance matrix of T and let $D_{ij} = d_{ij} - d_{in}$ if

$i \neq j$, and $D_{ii} = -\lambda - d_{in}$.

Then $T_1$ and $T_2$ are cospectral iff the determinants of

$$\begin{pmatrix} 0 & d_{13}-d_{23} & d_{14}-d_{24} & \cdots & d_{1n}-d_{2n} \\ d_{13}+d_{23} & -\lambda & d_{34} & \cdots & d_{3n} \\ d_{14}+d_{24} & d_{34} & -\lambda & & d_{4n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ d_{1n}+d_{2n} & d_{3n} & d_{4n} & \cdots & -\lambda \end{pmatrix} = L_1$$

and

$$\begin{bmatrix} D_{11}+D_{12} & D_{13} & D_{14} & \cdots & D_{1n} \\ D_{12}+D_{22} & D_{23} & D_{24} & \cdots & D_{2n} \\ \vdots & \vdots & \vdots & & \vdots \\ D_{n-1\,1}+D_{n-1\,2} & D_{n-1\,3} & D_{n-1\,4} & \cdots & D_{n-1\,n} \end{bmatrix} = L_2$$

are zero. Note that this theorem does not distinguish isomorphic cospectral pairs.

**Proof of Theorem 5** Theorem 4 gives us the contribution of a leaf $x_1$. Let $S=D(T)-\lambda I$. Since S is symmetric, $\det(S(j,i))=\det(S(i,j))$. Using the notation of the proof of Theorem 4, we have

$$CP(\Delta(T_1)) - CP(\Delta(T_2)) = -\lambda^2 \left(\det(S(1,1)) - \det(S(2,2))\right)$$

$$-2x_1\left( \lambda\det(S(1,1)) + \lambda\sum_{j=3}^{n}(-1)^j\det\left(S(1,j)\right) - \lambda\sum_{j=2}^{n}(-1)^j\det\left(S(2,j)\right) \right)$$

Now $\det(S(1,1)) = \det(S(2,2))$ iff $\det(L_1) = 0$. This is true because if we expand $\det(L_1)$ by diagonals, taking an entry from the first row and an entry from the first column, we get

$(-1)^{i+j-1}(d_{1i}-d_{2i})(d_{1j}+d_{2j})\det(S(1,1)(2,2)(j,i))$ and its symmetric

partner

$(-1)^{i+j-1}(d_{1j}-d_{2j})(d_{1i}+d_{2i})\det(S(1,1)(2,2)(i,j))$. Adding these

together we get

$(-1)^{i+j-1}2(d_{1i}d_{1j}-d_{2i}d_{2j})\det(S(1,1)(2,2)(i,j))$. But in the expansion

of $\det(S(1,\bar{1}))$, we get

$$(-1)^{i+j-1}d_{2i}d_{j2}\det(S(1,1)(2,i)(j,2)) + d_{i2}d_{2j}\det(S(1,1)(i,2)(2,j2)).$$

Similarly, we can work out the expression for $\det(S(2,2))$. This proves the equivalence. Showing that

$$\lambda\det(S(1,1)) + \lambda\sum_{j=3}^{n}(-1)^{j}\det\left(S(1,j)\right) = \lambda\sum_{j=2}^{n}(-1)^{j}\det\left(S(2,j)\right)$$ is

equivalent to $\det(L_2)=0$ is similar and not difficult.

## Section III

For any graph G, the following theorem of Graham, Hoffman and Hosoya breaks down the determinant of the distance matrix into a sum of terms related to the 2-connected components of the tree [GHH]. This theorem provides a most elegant proof that the determinant of the distance matrix of any tree on n vertices is the same. One simply observes that the 2-connected pieces of a tree are its edges.

**Theorem 6** [GHH] Let G be a finite graph in which each edge e has associated with it an arbitrary non-negative length w(e). (The usual weight chosen for edges is w(e)=1.) Let $d_{ij} = \min_{P(v_i,v_j)} w(P(v_i,v_j))$

where $P(v_i,v_j)$ ranges over all paths from $v_i$ to $v_j$ and $w(P(v_i,v_j))$

denotes the sum of all edge-lengths in $P(v_i,v_j)$. Let the 2-connected

pieces of G be $G_1,G_2,\dots,G_k$. Let $D(G)^{-1} = (d_{ij}^{-1})$ and cofactor(D(G))

$= \text{cof}(D(G)) = \det(D(G)) \sum_{i,j} d_{ij}^{-1}$ .

Then $\det(D(G)) = \sum_{i=1}^{k} \det(D(G_i)) \prod_{j\neq i} \text{cof}(D(G_j))$.

Recall that we promised a proof of Lemma 6 from the proof of Theorem 2, that the determinant of the distance matrix of a weighted

tree is $(-1)^{n-1}2^{n-2}\prod_{i=1}^{n-1} x_i \left( \sum_{i=1}^{n-1} x_i \right)$. Using Theorem 6, we can provide

an easy proof by observing that the cofactor of an edge weighted with

$x_i$ is $-2x_i$ and the determinant is $x_i^2$.

A simple lemma from [GHH] will aid us in computing the cofactor of a matrix. Recall that for any matrix M, M(i,j) is M with its ith row and $j^{th}$ column deleted. Let $M^*$ be M with the first row subtracted from every other row and the first column subtracted from every other column ($M^*$ retains the first row and column of M).

<u>Lemma 10</u> [GHH] For any matrix M, $cof(M) = det(M^*(1,1))$.

A version of Theorem 6 that includes the other coefficients of the characteristic polynomial would seem difficult, as demonstrated by the straightforward attempt below.

**Theorem 7** Let G = $G_1$ and $G_2$ joined together at vertex v.
Let $D(G_i) - \lambda I = M_i$ for i=1,2. Then

$$CP(\Delta((G)) = det(M_1)cof(M_2) + det(M_2)cof(M_1)$$

$$+ \lambda\Big(cof(M_1^*)cof(M_2) + det(M_1)cof(M_2^*(1,1)) + cof(M_2^*)cof(M_1)$$

$$+ det(M_2)cof(M_2^*(1,1))\Big)$$

$$+ \lambda^2\Big( cof(M_1^*)cof(M_2^*(1,1)) + cof(M_2^*)cof(M_1^*(1,1))\Big)$$

**Proof of Theorem 7** The proof involves a straightforward application of the techniques of the proof of Theorem 6 and is not enlightening. Hence I do not include it here.

We have a description of the coefficients of CP(D(P)) where P is a path using Theorem 6. Let $D\{v_1,v_2,...,v_k\}$ be the kxk submatrix of D(T) whose rows and columns are indexed by $v_1,v_2,...,v_k$. Then the coefficient of $\lambda^{n-k}$ in CP(D(T)) is (for $k \geq 2$)

$$(-1)^{n-k} \sum_{1 \leq i_1 < i_2 < \cdots < i_k \leq n} \det\left(D\{v_{i_1}, v_{i_2}, \ldots, v_{i_k}\}\right).$$

When T is a path P, we may interpret $D\{v_{i_1}, v_{i_2}, \ldots, v_{i_k}\}$ as the distance matrix of the path $P\{i_1, i_2, \ldots, i_k\}$ whose vertices are $v_{i_1}, v_{i_2}, \ldots, v_{i_k}$ and whose edges are the k-1 shortest distances from the set of distances between $v_{i_1}, v_{i_2}, \ldots, v_{i_k}$. Since we started with a path, $P\{i_1, i_2, \ldots, i_k\}$ is still a path and we can apply Theorem 6 above to get

**Theorem 8** Let P be a path with n vertices. The coefficient of $\lambda^{n-k}$ in CP(D(P)) is

$$(-1)^{n-1} 2^{k-2} \sum_{1 \leq s_1 < s_2 < \cdots < s_k \leq n} (s_2-s_1)(s_3-s_2)\cdots(s_k-s_{k-1})(s_k-s_1)$$

**Proof of Theorem 8** Since $P\{i_1, i_2, \ldots, i_k\}$ is a tree, its 2-connected pieces are its edges. Therefore, its determinant is

$$-(-2)^{k-2}(s_2-s_1)(s_3-s_2)\cdots(s_k-s_{k-1})[(s_2-s_1)+(s_3-s_2)+\cdots+(s_k-s_{k-1})]$$

since the distance between $s_2$ and $s_1$ is $(s_2-s_1)$. When we multiply this by $(-1)^{n-k}$ we get the above.

Theorem 8 could possibly be used to show the unimodality of the coefficients of a path as described in the following conjecture.

**Conjecture 1** The sequence $(1/2^{n-2})|d_0(T)|$, $(1/2^{n-3})|d_1(T)|$, ... is unimodal, with the peak occuring at $(1/2^{n-[n/2]-2})|d_{[n/2]}(T)|$ (due to R. L. Graham).

We know exactly how to describe $d_0$ and we have a closed form expression for $d_1(T)$ from Theorem 2. However, we can also look at $d_1(T)$ from this point of view:

$$d_1(T) = -\sum_{i=1}^{n} \det(D(T)(i,i)).$$

Recall that $D(T)(i,i)$ is the distance matrix of the graph $G_i$ with the vertex $v_i$ in T replaced with the complete graph $K_{deg(v_i)}$ on the neighbors of $v_i$. The edge between $v_j$ and $v_k$, neighbors of $v_i$, is now weighted 2 in the distance matrix. We can apply Theorem 6 to each $G_i$ and find the determinant $\det(D(G_i))$ by computing the determinant and cofactor of each of its 2-connected components. These are its complete graph $K_{deg(v_i)}$ and the remaining edges that come directly from T.

**Remark** The coefficient of $\lambda$ in $CP(D(T))$ is

$$(-1)^{n-1}2^{n-3}\sum_{v \text{ in } T}(-deg(v)^2 + deg(v)(n+3) - 4).$$

**Proof of Remark** Let $deg(v_i) = g_i$. Now $d_1(T) = -\sum_{i=1}^{n} \det(D(T)(i,i))$.

Using Theorem 6, we can write $\det(D(T)(i,i)) = (2^{g_i}\det(K_{g_i}))(-2)^{n-g_i-1}$ $+ (n-1-g_i)(-1)(-2)^{n-g_i-2}(2^{g_i-1}\mathrm{cof}(K_1))$. Lemmas 11 and 12 below complete the proof.

<u>Lemma 11</u> $\det(D(K_i)) = (-1)^{i-1}(i-1)$

<u>Lemma 12</u> $\mathrm{cof}(D(K_i)) = (-1)^{i-1} i$

Proof of Lemma 12: $$D(K_i)^{-1} = \frac{1}{i-1}\begin{bmatrix} -i+2 & 1 & 1 & & 1 \\ 1 & -i+2 & & & 1 \\ & & \cdot & \cdot & \cdot \\ 1 & 1 & & & -i+2 \end{bmatrix}$$

Hence $$\mathrm{cof}(K_i) = (-1)^{i-1}(i-1)\left(\frac{1}{i-1}\left(i(-i+2) + i^2 - i\right)\right) = (-1)^{i-1} i.$$

This remark hasn't given us much new information. However, we can use the same technique to get a weighted version of the above. It seems worthwhile to compute the cof and det of a weighted complete graph, but the coefficient of $\lambda$ that we obtain in this way is not very enlightening.

<u>Lemma 13</u> Let H be a graph on vertices $\{1,2,\ldots,i\}$, composed of disjoint cycles of size $\geq 2$ (i.e., a 2-cycle has two copies of the same edge and hence its total weight is the square of the weight of the edge) and with no isolated vertices. The weight of edge e between vertices k and j is the variable $y_{jk}$. Let f = the number of even cycles of H. Let f' = the number of cycles of size $\geq 3$ in H.

Then the determinant of $$Y = \begin{bmatrix} 0 & y_{12} & y_{13} & \cdots & y_{1i} \\ y_{12} & 0 & y_{23} & \cdots & y_{2i} \\ \cdot & & \cdot & & \\ y_{1i} & y_{2i} & y_{3i} & \cdots & 0 \end{bmatrix} \text{ is } \sum_H \omega(H)$$

where $\omega(H) = (-1)^{f} 2^{f'} \prod_{e \subset_r H} w(e)$.

Proof of Lemma 13 The proof is simply expanding the determinant of Y as

$\sum_{\sigma \in S_i} (-1)^{\mathrm{sign}(\sigma)} y_{1\sigma(1)} y_{2\sigma(2)} \cdots y_{i\sigma(i)}$ where $\sigma$ is a member of $S_i$, the permutation group on i letters, with $y_{jj}$ taken to be zero. The power of 2 comes from observing that each cycle can appear twice due to the symmetry of Y.

Lemma 14 Let H* be a graph on vertices {1,2,...,i}, composed of a disjoint collection of a path with at least one vertex and some number of cycles of size >2, possibly none. The weight of edge e between vertices k and j is the variable $y_{jk}$. Let f* = i + #-connected components of H*. Let f** = #-cycles of size>3 + 1 - #-isolated vertices.

$$\text{Then the cofactor of } Y = \begin{bmatrix} 0 & y_{12} & y_{13} & \cdots & y_{1i} \\ y_{12} & 0 & y_{23} & \cdots & y_{2i} \\ \cdot & & \cdot & & \\ y_{1i} & y_{2i} & y_{3i} & \cdots & 0 \end{bmatrix} \text{ is } \sum_{H^*} \psi(H^*)$$

where $\psi(H^*) = (-1)^{f^*} 2^{f^{**}} \prod_{e \subset_r H^*} w(e)$.

Proof of Lemma 14 Notice that in the definition of H* above a 2 vertex subset of {1,2,...,i} can appear as a path with weight the weight of the edge between the vertices, or as a cycle with the square of the weight of the edge between the vertices. We use Lemma 5 from above to write cof(Y) as a determinant:

$\mathrm{cof}(Y) = \det(Y_{jk})\Big|_{j,k=2}^{i}$ where $Y_{jk} = y_{jk} - y_{1j} - y_{1k}$ and $y_{jj} = 0$. When we expand the determinant as a sum of diagonals, as in Lemma 9, $\sum_{\sigma\in S_i} (-1)^{\mathrm{sign}(\sigma)} Y_{2\sigma(2)} Y_{3\sigma(3)} \cdots Y_{i\sigma(i)}$ we observe that each of the terms is a polynomial in the $y_{jk}$'s with degree i-1. Each member of the set {2,3,...,i} appears at most twice as a subscript in any monomial from the terms above. Since there was nothing special about expanding the cofactor using the first row and column, we can conclude that det(Y) is a sum of monomials in the $y_{jk}$ that satisfy that in any monomial, no subscript appears more than twice. Take each monomial as representing the product of the weights of the edges of the graph $G_m$ with the i-1 edges of $G_m$ given by its factors. Then the vertex degrees in $G_m$ are all less than or equal to two and the graph has i-1 edges. Hence it is a union of cycles and paths, and since there are i-1 edges, there is exactly one path.

The sign of a monomial is $(-1)^{\mathrm{sign}(\sigma)} + \deg(1)$ in $G_m$, where $\sigma$ represents the ordering the the $Y_{jk}$'s given above. Since $G_m$ always contains a path, without loss of generality, we may assume deg(1) = 0 or 1. If deg(1)=0, then sign($\sigma$) = the # of even cycles in $G_m$, and hence is congruent to i + #-connected components of $G_m$. If deg(1) = 1 then sign($\sigma$) = the # of even cycles of $G_m$ + the # of odd paths in $G_m$, so 1 + sign($\sigma$) is again congruent to i + the # of connected components of $G_m$. In order to get the power of 2 in the statement of the lemma, we observe that each cycle and path can occur twice because of the symmetry of Y. However, the path containing a single vertex does not contribute a power of 2.

**Theorem 9** Let $K\Delta_j$ = the weighted distance matrix of the complete graph on the neighbors of $v_j$ with the weight of the edge between $v_r$ and $v_s$ as $x_{jr} + x_{js}$. Let $G\Delta_j$ be the complete graph whose distance matrix is $K\Delta_j$. Let $g_i = \deg(v_i)$. If an edge e is not adjacent to a vertex $v_j$, we say $e \not\in v_j$.

The coefficient of $\lambda$ in $CP(\Delta(T))$ is

$$= \sum_{v_i \text{ in } T} (-1)^{n-1-g_i} 2^{n-2-g_i} \prod_{x_j \not\in v_i} x_j \left( 2 \sum_{H \subset G\Delta_i} r\, \omega(H) + \sum_{H^* \subset G\Delta_i} r \psi(H^*) \left( \sum_{x_j \not\in v_i} x_j \right) \right)$$

**Proof of Theorem 9**

Now $\delta_1(T) = -\sum_{i=1}^{n} \det(\Delta(T)(i,i))$. Using Theorem 6, we can write

$$\det(\Delta(T)(i,i)) = \det(K\Delta_i) \prod_{x_j \not\in v_i} \operatorname{cof}\begin{pmatrix} 0 & x_j \\ x_j & 0 \end{pmatrix}$$

$$+ \sum_{x_j \not\in v_i} \operatorname{cof}(K\Delta_i) \det\begin{pmatrix} 0 & x_j \\ x_j & 0 \end{pmatrix} \prod_{\substack{x_k \not\in v_i \\ k \neq j}} \operatorname{cof}\begin{pmatrix} 0 & x_j \\ x_j & 0 \end{pmatrix}$$

$$= \det(K\Delta_i)(-2)^{n-1-g_i} \prod_{x_j \not\in v_i} x_j$$

$$+ \operatorname{cof}(K\Delta_i)(-1)^{n-1-g_i}2^{n-2-g_i}\left(\sum_{x_j \notin v_i} x_j\right)\left(\prod_{x_j \notin v_i} x_j\right)$$

$$= (-1)^{n-1-g_i}2^{n-2-g_i}\prod_{x_j \notin v_i} x_j \left(2\det(K\Delta_i) + \operatorname{cof}(K\Delta_i)\left(\sum_{x_j \notin v_i} x_j\right)\right)$$

Hence the coefficient of $\lambda$ in $CP(\Delta(T))$ is

$$\sum_{v_i \text{ in } T} (-1)^{n-1-g_i}2^{n-2-g_i}\prod_{x_j \notin v_i} x_j \left(2\det(K\Delta_i) + \operatorname{cof}(K\Delta_i)\left(\sum_{x_j \notin v_i} x_j\right)\right)$$

Plugging in Lemmas 13 and 14, we get the Theorem 9.

## Section IV

We now proceed to factor CP(D(T)) using the automorphism group of the tree. The method involved is to find subspaces V of D(T) such that $D(T)\cdot V \subseteq V$. We then compute the characteristic polynomial of the action of D(T) on the subspaces. This must divide the characteristic polynomial of the distance matrix of the whole tree. Let $e_i$ be the vector with 1 in the ith place and 0 otherwise. For any set S, let $e_S = \sum_{s\in S} e_s$ .

**Theorem 10** For any graph G, we always have a factor of degree the number of orbits of G. The associated vector space is defined in the proof below. The factor from this vector space is called the orbit factor.

**Proof of Theorem 10** Let the vertex orbits of G be $S_1, S_2, \ldots , S_r$. Let $V_{orbit} = \langle e_{S_i} : 1 \leq i \leq r\rangle$. Then row(j) times $e_{S_i} = \sum_{s\in S_i} d_{j,s}$ . Since every automorphism of G preserves distance, $\sum_{s\in S_i} d_{j,s} = \sum_{s\in S_i} d_{j',s}$ for every j' in the same orbit as j. This completes the proof.

**Theorem 11** Suppose T has a subgraph with a vertex v joined to two isomorphic copies of rooted tree $T^*$ at root u. Suppose that v disconnects each copy of $T^*$ from the rest of T. Then we can find a factor of CP(D(T)) of degree the number of orbits of rooted tree $T^*$.

**Proof of Theorem 11** Let the left-sided copy of $T^*$ be labelled $T^*$ with root u and the right-sided copy of $T^*$ be labelled $\mathcal{T}^*$ with root $\mathcal{u}$. Let the vertex orbits of rooted tree $T^*$ be $S_1, S_2, \ldots, S_q$ and the orbits of $\mathcal{T}^*$ be $\mathcal{s}_1, \mathcal{s}_2, \ldots, \mathcal{s}_q$ with the correspondence $S_i \rightarrow \mathcal{s}_i$ in the

isomorphism between $T^*$ and $\mathcal{T}^*$. Let $V = \langle e_{S_i} - e_{\mathcal{S}_i} : 1 \leq i \leq q \rangle$.
We show that $D(T) \cdot V \subseteq V$. It is clear that the dimension of $V$ is the number of orbits of rooted tree $T^*$.

Now $D(T) \cdot (e_{S_i} - e_{\mathcal{S}_i}) = \left( \sum_{s \in S_i} d_{1s} - \sum_{\mathscr{s} \in \mathcal{S}_i} d_{1\mathscr{s}}, \ldots, \sum_{s \in S_i} d_{ns} - \sum_{\mathscr{s} \in \mathcal{S}_i} d_{n\mathscr{s}} \right)$

If $w$ is in $T$, but not in $T^*$ or $\mathcal{T}^*$, then $\sum_{s \in S_i} d_{ws} - \sum_{\mathscr{s} \in \mathcal{S}_i} d_{w\mathscr{s}} = 0$,

since $d_{ws} - d_{w\mathscr{s}} = (d_{wu} + d_{us}) - (d_{w\mathscr{u}} + d_{\mathscr{u}\mathscr{s}}) = 0$, because $s$ and $\mathscr{s}$ are in

the "same" orbit of $T^*$. If $w$ is in $T$ and $\mathscr{w}$ is its image under the isomorphism, then $\sum_{s \in S_i} d_{ws} - \sum_{\mathscr{s} \in \mathcal{S}_i} d_{w\mathscr{s}} = -\left( \sum_{s \in S_i} d_{\mathscr{w}s} - \sum_{\mathscr{s} \in \mathcal{S}_i} d_{\mathscr{w}\mathscr{s}} \right)$ by symmetry.

If $w_1$ and $w_2$ are in the same orbit in $T^*$, then $\sum_{\mathscr{s} \in \mathcal{S}_i} d_{w_1\mathscr{s}} = \sum_{\mathscr{s} \in \mathcal{S}_i} d_{w_2\mathscr{s}}$ since

$w_1$ and $w_2$ are the same distance away from $u$ and must go through $u$ to get to $\mathcal{T}^*$. Now $\sum_{s \in S_i} d_{w_1 s} = \sum_{s \in S_i} d_{w_2 s}$ since the isomorphism that makes

$w_1$ and $w_2$ in the same orbit must preserve distance. Hence the action of $D(T)$ on $V$ preserves $V$.

We use Theorem 11 to compute $CPD(B_r)$, where $B_r$ is the full binary tree with maximum length $r$ from the top vertex, by finding linearly independent vector spaces that are preserved under the action of $D(T)$. For each vertex not on the bottom level we have exactly the situation of the preceding theorem. Let the top level be labelled 0 and the level at distance $i$ from the top be labelled level $i$. Each vertex on level $i$ forms a $B_{r-i}$ with all its descendants. This generates a factor of $CP(D(B_r))$ of degree $r-i$. Each factor of degree $r-i$ appears in

$CPD(B_r)$ $2^i$ times, since level i contains $2^i$ vertices. We also get a factor of degree r+1 by taking $V_{orbit}$ of the whole tree.

**Theorem 12** Let $B_r$ be the full binary tree with maximum distance r from the top vertex. Then $CP(D(B_r)) = P_1^{2^{r-1}} \cdot P_2^{2^{r-2}} \cdots P_r \cdot Q_{r+1}$ where each $P_i$ has degree i and in fact, $P_{i+1}(\lambda) = -(2+3\lambda)P_i(\lambda) + 2\lambda^2 P_{i-1}(\lambda)$. Each $P_i$ has leading coefficient $(-1)^i$ and $P_1 = -(\lambda+2)$, $P_2 = \lambda^2 + 8\lambda + 4$.

**Proof of Theorem 12** We first show that the vector spaces associated with the polynomials from Theorem 11 are linearly independent. The vertex orbits of $B_r$ are the levels of the tree. All of the vector spaces given by Theorem 11 are generated by vectors with non-zero entries only in one orbit. Hence we need only consider a single orbit at a time to determine if all the vector spaces are independent. Let $S(m) = \{1,2,\ldots,2^m\}$ and let $M(p)$ be the $2^p x 2^p$ matrix whose rows are given by $e_{S(p)}$, $e_{S(p-1)} - e_{2^{p-1}+S(p-1)}$, $e_{S(p-2)} - e_{2^{p-2}+S(p-2)}$,

$e_{2\cdot 2^{p-2}+S(p-2)} - e_{3\cdot 2^{p-2}+S(p-2)}, \ldots, e_1 - e_2, e_3 - e_4, \ldots, e_{2^p-1} - e_{2^p}$.

It is easy to show that det M(p) is never zero. We replace $e_{S(p-1)} - e_{2^{p-1}+S(p-1)}$ with itself plus $e_{S(p)}$ to get $2e_{S(p-1)}$ and then replace $e_{S(p)}$ with itself minus $e_{S(p-1)}$ to get $e_{2^{p-1}+S(p-1)}$. We then have that $|\det M(p)| = 2(\det M(p-1))^2$. When p = 1, we get that $\det M(1) = -2$. Hence by induction det M(p) is never zero. Therefore the vector spaces are independent.

Using Theorem 10 and 11, we now have shown Theorem 12, except for the recursion of the P's. To do this we must compute the matrix of the action of $D(B_r)$ on the vector spaces of Theorem 12. Fix r and let M be the $(2^{r+1} - 1)\mathrm{x}r$ matrix with columns $\{e_{S_i} - e_{\mathcal{S}_i} : 1 \le i \le q\}$ where $S_i$ is the set of vertices in $B_r$ at distance i from the top and descendants of the left-hand child of the top, and $\mathcal{S}_i$ is the other half of vertices on level i.

Now $P_r$ is the characteristic polynomial of the rxr matrix Π that records the action of $D(B_r)$ on M. That is, the $ij^{th}$ entry of matrix Π is the coefficient of $e_{S_i} - e_{\mathcal{S}_i}$ in the linear expansion of $D(B_r)\cdot(e_{S_i} - e_{\mathcal{S}_i})$. Since we have one vector $e_{S_i} - e_{\mathcal{S}_i}$ for each orbit $S_i$, we need compute only $\mathrm{row}(j)\cdot(e_{S_i} - e_{\mathcal{S}_i})$ for a single vertex j in orbit $S_j$ to find the $ij^{th}$ entry of Π.

If $i \le j$, then row(j) in $D(B_r)$ times $e_{S_i} - e_{\mathcal{S}_i}$ is

$$(j-i) + (j-i+2) + 2(j-i+4) + 4(j-i+6) + \ldots + 2^{i-2}(i+j-2) - 2^{i-1}(i+j)$$

$$= -2(2^{i-1} - 1).$$

If $i > j$, then row(j) in $D(B_r)$ times $e_{S_i} - e_{\mathcal{S}_i}$ is

$$2^{i-j}\Big( (i-j) + (i-j+2) + 2(i-j+4) + \ldots + 2^{j-2}(i+j-2) - 2^{j-1}(i+j)\Big)$$

$$= -2^{i-j+1}(2^{j-1} - 1).$$

So $P_r = \det(\Pi - \lambda I)$. We perform some elementary row and column operations on $\Pi-\lambda I$ to simplify taking its determinant. First we divide every row by −2 and substitute γ for λ/2. Then we subtract the $k^{th}$ column from the $k+1^{st}$ column, beginning with k = r (the first column remains the same). Finally we subtract 2 times new row k from new row k+1 for k = 1 to r. Our resulting matrix $Z_r = (z_{ij})$ has $z_{ij} = 0$ if $|i-j| \geq 2$, $z_{11} = 1+\gamma$, $z_{ii} = 1+3\gamma$ for $i>1$, $z_{i,i+1} = -\gamma$ and $z_{i+1,i} = 2\gamma$. By expanding out the last column, we get that

$\det(Z_r) = (1+3\gamma)\det(Z_{r-1}) + 2\gamma^2\det(Z_{r-2})$.

When we multiply by $(-2)^r$ and substitute $\lambda/2$ for $\gamma$, we get

$P_r(\lambda) = -(2+3\lambda)P_{r-1}(\lambda) + 2\lambda^2 P_{r-2}(\lambda)$.

We mention without proof that Theorem 12 can be generalized for full k-nary trees. Let $F_r$ be the full k-nary tree with maximum distance r from the top vertex. Then CPD($F_r$) = $P_1^{(k-1)k^{r-1}} \cdot P_2^{(k-1)k^{r-2}} \cdots P_r \cdot Q_{r+1}$ where each $P_i$ has degree i and $P_{i+1}(\lambda) = -(2+(k+1)\lambda)P_i(\lambda) + k\lambda^2 P_{i-1}(\lambda)$. The proof of this will appear in an upcoming paper [C]. There will also appear in [C] a more complicated recursion developed for the $Q_r$'s, which are the polynomials associated with the $V_{orbit}$ of the full binary tree $B_r$.

Another attempt to find subspaces to use to factor the CP of the distance matrix led to the following.

**Theorem 13** Let s be an involution of the automorphism group of T, contained in $S_n$. Let $V_s$ be generated by $\{e_1 - e_{s(1)}, e_2 - e_{s(2)}, \ldots\}$. Then the action of D(T) on $V_s$ preserves $V_s$, that is, $D(T)\cdot V_s \subseteq V_s$.

**Proof of Theorem 13**

$D(T)\cdot(e_i - e_{s(i)}) = [d_{1i} - d_{1s(i)}, d_{2i} - d_{s(i)}, \ldots, d_{ni} - d_{ns(i)}]$

But $d_{1i} - d_{1s(i)} = -(d_{s(1)i} - d_{s(1)s(i)})$ since s is an automorphism of T and therefore preserves distance and also since s is an involution. Hence $D(T)\cdot V_s$ is contained in $V_s$.

Using Theorems 10 and 13, we can factor the CP of the distance matrix of a path into "halves", much like the effect of its single automorphism, the flip. We get a factor of degree $\lfloor n/2 \rfloor$ from its non-trivial involution. We also get an orbit factor of degree $\lceil n/2 \rceil$. Thus we can write the CP of a path as the product of two factors, each of degree approximately half of the number of vertices.

Theorem 13 also allows us to factor the characteristic polynomial of the distance matrix of a graph, not just a tree. For example, we can factor the CP of the odd cycle $C_{2k+1}$ by taking two involutions, say, the flip that leaves vertex 1 fixed and the flip that leaves vertex 2 fixed plus the factor from $V_{orbit}$, which is $(\lambda-(n^2+n))$. Each flip gives us a factor of degree n. We can easily factor the CP of the distance matrix of a complete graph on n vertices minus an edge: we get an orbit factor of degree 2, an involution factor that switches the vertices of degree n-2, and n-3 involutions that switch the vertices of degree n-1 to get $(-1)^n(\lambda^2 - (n-1)\lambda - 2)(\lambda+2)(\lambda+1)^{n-3}$.

**Group Representation Connection**

People knowledgeable about representation theory will tell you that the results in this section are no surprise. Here is a brief description of why the automorphism group helps factor the characteristic polynomial. The group representation backgound can be found in [I]. If we take the representation of the automorphism group of G as the permutation representation on the vertices of G, realized by nxn matrices, then this representation has a character $\psi$. This character $\psi$ can be written as the sum of irreducible characters, and corresponding to each irreducible character $\chi$ there is a vector space $V_\chi$, called the isotypic component of $\chi$, that is invariant under the action of the automorphism group of G. If $\chi$ appears with multiplicity greater than 1 in the linear expansion of $\psi$, then $V_\chi$ breaks up into a sum of vector spaces, $V_{\chi_1}$, $V_{\chi_2}$, ..., $V_{\chi_{m(\chi)}}$, (with $m(\chi)$ = the multiplicity of $\chi$) each of which are invariant under the automorphism group of G. The distance matrix of G is fixed by the automorphism group of G and hence, by Schur's Lemma, D(G) acts as a scalar $\alpha_{\chi_i}$ on each $V_{\chi_i}$. Then what we have observed is that

$$CP(D(G)) = \prod_{\chi \text{ appears in the linear expansion of } \psi} \left\{ (\lambda-\alpha_{\chi_1})(\lambda-\alpha_{\chi_2}) \ldots (\lambda-\alpha_{\chi_{m(\chi)}}) \right\}^{\dim(V_{\chi_1})}$$

## CONJECTURES

**Conjecture 1** The sequence $(1/2^{n-2})|d_0(T)|$, $(1/2^{n-3})|d_1(T)|$, ... is unimodal, with the peak occuring at $(1/2^{n-[n/2]-2})|d_{[n/2]}(T)|$ (due to R. L. Graham).

**Conjecture 2** Binary trees are characterized by the characteristic polynomial of their distance matrices. (due to Lynne Butler)

**Conjecture 3** If two trees have the same CP(D(T)), they have the same degree sequence. (due to M. A. Hovey)

**Open Problem 4** What other classes of trees have easily factorable CP(D(T))'s?

**Conjecture 5** If two trees have the same CP(D(T)), then they have the same automorphism group.

**Conjecture 6** CP(D(T)) of a tree with trivial automorphism group has no rational roots.

**References**

[A] Sabra S. Anderson, Graph Theory and Finite Combinatorics, Markham Publishing Co., Chicago, 1970.

[C] K. L. Collins, "Factoring the Characteristic Polynomial of Distance Matrices" to appear.

[CDS] Dragos M. Cvetkovic, Michael Doob, Horst Sachs, Spectra of Graphs, Academic Press, New York, 1980, vol 87 in the series "Pure and Applied Math".

[GHH] R. L. Graham, A. J. Hoffman, and H. Hosoya, "On the Distance Matrix of a Directed Graph", **Journal of Graph Theory,** Vol I, 85-88, 1977.

[G&L] R. L. Graham and L. Lovász, "Distance Matrix Polynomials of Trees" **Advances in Math.,** Vol 29, 60-88, July 1978.

[I] I. Martin Isaacs, Character Theory of Finite Groups, Academic Press, New York, 1976.

[McK] B. D. McKay, "On the spectral characterization of trees", **Ars Combinatoria III** (1977), 219-232.

**List of Notation**

$a_i$ = the coefficient of $\lambda^i$ in CP(A(T)).

$a_{ij}$ = 1 if vertex $v_i$ and $v_j$ are adjacent; 0 otherwise.

$A_F^k$ = the integer coefficient in Theorem 1 that does not depend on T.

$\mathcal{A}_F^k$ = the polynomial coefficient in Theorem 2 that corresponds to $A_F^k$.

A(T) = the adjacency matrix of tree T.

A(G) = the adjacency matrix of graph G.

$\alpha_{ij} = a_{ij} / x_{ij}$

B = the matrix that makes $\Delta(T) = N^T B N$ true.

$|C|$ = the number of vertices in component C.

$||C||$ = the number of edges in component C.

cofactor(M) = cof(M) = $\det(M) \sum m_{ij}^{-1}$ for matrix M.

CP(M) = the characteristic polynomial of matrix M.

$d_i$ = the coefficient of $\lambda^i$ in CP(D(T)).

$d_{ij}$ = the distance between vertex $v_i$ and vertex $v_j$.

d(u,v) = the distance between u and v.

D(T) = the distance matrix of tree T.

D(G) = the distance matrix of graph G.

$\delta_i$ = the coefficient of $\lambda^i$ in CP($\Delta$(T)).

$\delta_{ij}$ = the weighted distance between vertex $v_i$ and vertex $v_j$.

$\delta$(u,v) = the weighted distance between u and v.

$\Delta$(T) = the weighted distance matrix of tree T.

$\Delta$(G) = the weighted distance matrix of graph G.

$\Delta^{-1}$ = the inverse of $\Delta$(T).

e = special edge labelling in the proof of Theorem 2.

$|F|$ = the number of vertices in forest F.

$||F||$ = the number of edges in forest F.

$g_i$ = degree of vertex $v_i$.

$G_i$ = a 2-connected component of graph G.

$L(i)$ = the number of i-matchings (in an unspecified graph).

M = usually denotes a matrix.

$M^*$ = the matrix M with its first row subtracted from every other row and its first column subtracted from every other column.

$N_F(T)$ = the number of copies of forest F contained as a subforest of tree T.

$N = (n_{ij})$ = the matrix given by $n_{ij} = 1$ if j is on the unique path from 1 to i and 0 otherwise.

$\nu_{ij}$ = the $ij^{th}$ entry of $N^{-1}$.

$P_k$ = the path on k edges.

$P(u,v)$ = the path from u to v.

$\rho(F) = \prod |C|$ where C is a component of F.

T = a tree.

$v_i$ = the $i^{th}$ vertex of tree T.

V = a vector space.

$x_i$ = a variable label on the $i^{th}$ edge of tree T.

$x_{ij}$ = the variable label of the edge between vertex $v_i$ and $v_j$.

X = the sum of all the variable labels on the edges of weighted tree T.

**A Philosophical Conclusion**

Most mathematicians love to play games. In fact, they love it so much that they do it all the time. Even more characteristically, mathematicians love to play games with changing rules, and would claim that the fun lay in deciding which rules to choose instead of following the ones already picked out. Unfortunately, all this game playing can be taken to an extreme that, it seems to me, destroys the fun of the game.

As we all know, the number of mathematicians in the world has been increasing dramatically, as has the body of mathematical knowledge. Many journals have sprung up and it is not easy to keep track of what's going on in one's own field, much less what's going on in other fields. More than that, the standards of "good" mathematics have become more and more blurred as an inrush of material to be read, digested or understood and graded--i.e. by accepting for publication--has flowed in another tenfold. No one has the time to read yet another new and probably not very interesting--that is, unlikely to directly relate to my theorem on characteristic polynomials of the distance matrices of binary trees--paper written by yet another underpaid, overworked mathematician struggling to meet the standards of expertise that are never written down, and not well understood.

Graduate schools, whose job one would believe to be that of setting the standards of graduating mathematicians, concentrate on a single aspect of a mathematician's life to the exclusion of the rest of mathematical experience. Admittedly, we all know that the mathematicians who have jobs at graduate schools have not been taught how to teach others, have not been taught how to give advice, have not been taught to describe the process of mind that goes on behind mathematics, so the best they can do is to encourage the students to learn about these things the way they did--an ad hoc procedure developed by trial and error--a procedure that has been used for hundreds of years by mathematicians without being improved. We like to think of mathematical methods as always being improved, but we have not even begun on the basics yet.

An unfortunate effect of our emphasis on newness instead of quality, is that, instead of consolidating what we have, we are constanly struggling not to repeat the poorly written work of someone on a different campus. The requiste to graduate from the graduate program of a high quality graduate schools is "original research". If the student's work has been duplicated before graduation, that work doesn't "count". On the other hand, no matter how uninteresting the work of said student, as long as it hasn't been previously published, it's usually enough to get a degree. Students are not asked to write well--it's considered a bonus if they can. Certainly no one may expect that after graduation the newly appointed professor will include writing practice as part of the professorial duties. Other professorial duties are certain to interfere--classes, homework, tests--duties that already take too much time away from "research".

However, when asked to explain the work that requires such a high priority and so much time, or even to explain the value of such work, most mathematicians will say that they "can't". Not only can they not explain it to others, a lot of the time they can't explain it to themselves. Now why is it that something they feel is so important doesn't receive enough of their attention for them to give a good description of it? The answer seems to be that we mathematicians haven't devoted enough time to thinking about what quality work is. It's usually easy for us to decide if a theorem is high quality--in retrospect--but standards for how to think about, how to describe, and how to write mathematics seem to be sorely lacking. Unfortunately, our confusion about writing mathematics can only lead to confusion about thinking mathematics.