THE CHARACTERISTIC POLYNOMIAL OF A GRAPH

by

Frank H. Clarke

Department of Mathematics
McGill University

Master in Science

A B S T R A C T

An expression for the characteristic polynomial of a graph is developed, showing the relationship between certain structural characteristics of the graph and the coefficients of its polynomial. Among other applications, a bipartite graph is shown to be characterized by its polynomial. A problem of Collatz is then investigated and solved for trees, and further results of the same nature are presented. A theorem on 1-factors in trees related to a theorem of Tutte is proven. It is shown that the polynomial of a graph yields certain information concerning coverings and line independence. In particular a formula for the point-covering number of a tree is established. The graph polynomial is then applied to problems related to Ulam's conjecture and graph reconstructions.

THE CHARACTERISTIC POLYNOMIAL OF A GRAPH

by

Frank H. Clarke

Clarke, F.H. Mathematics M.Sc.

Preface

The polynomial of a graph (as we will use the term) is a natural outgrowth of the concept of the adjacency matrix of a graph, which was defined in the pioneer work of König (11, p.237) in 1936. However, the first to actually investigate the properties of this polynomial were Collatz and Sinogowitz (1) in a paper published in 1957. Since then it has received more attention (as we shall see), but very little from the point of view of combinatorial properties of its coefficients.

Most of this thesis will concern itself with such properties. It is oriented toward obtaining new results rather than exposition of what has been discovered. The seven theorems and three propositions proven herein are original. The previous results of which we make use are of course credited in each case.

Theorem I is a fundamental characterization of the polynomial of a graph in terms of certain types of its subgraphs.

All the other theorems and propositions rely at least in part on this Theorem, and one could think of them as applications of it. We obtain results on bipartite graphs, structure of trees, coverings, and a problem suggested by Collatz (1). We also apply graph polynomials to Ulam's conjecture and reconstructions, an application which appears not to have been known previously.

I would like to thank Professor W. G. Brown for his advice in the completion of this work.

THE CHARACTERISTIC POLYNOMIAL OF A GRAPH

by

Frank H. Clarke

A thesis submitted to the Faculty of Graduate Studies and Research at McGill University, in partial fulfillment of the requirements for the degree of Master in Science.

Department of Mathematics
McGill University
Montreal

July 1970

TABLE OF CONTENTS

Continued...

Page

Page

I.1. This chapter is meant to serve as an introduction to the structures and concepts that we shall be using. Notation and terminology in graph theory have not been standarized to a great extent. We shall use mostly notation and terminology in accord with the recent book of Harary (3), indicating any new or uncommon definitions as they arise later on.

I.2. A *graph* G consists of a finite set V(G) of p *points* or *vertices* together with a set E(G) of unordered pairs of distinct points of V(G). Each (unordered) pair (u,v) of points in E(G) is a *line* or *edge* of G. We may also label this line by x = (u,v) and we say x *joins* u and v. The *null-graph* $\emptyset$ has no points and hence no lines.

Two points u and v of V(G) are *adjacent* iff[1] (u,v) is in E(G) (i.e. u and v are joined). Two different edges are *incident* if they have one common point, otherwise they are *disjoint*. An edge and a point are *incident* when the point is one of the two points making up that particular edge. The *valency* of a vertex is the number of edges incident to it. We will not allow *loops* (i.e. an edge from a vertex to itself).

Two graphs G and H are *isomorphic* ($G \simeq H$) if there exists a one-to-one correspondence f between their points such that fg_1 and fg_2 are adjacent points in H iff g_1 and g_2 are adjacent points in G.

---

(1) iff = if and only if

A *subgraph* H of G is a graph having all of its points and lines in G. We shall use (this definition is not standard) $|G|$ to mean the number of points in G, i.e. card (V(G)). A *spanning subgraph* S of G is a subgraph of G such that $|S| = |G|$. A graph G is called *odd* or *even* according to whether $|G|$ is odd or even.

When it is possible to partition the points of G into two nonempty classes such that a vertex from one class is never adjacent to a vertex from the other, we say G is *disconnected*; otherwise G is *connected.*

A maximal connected subgraph of G is called a *connected component* of G, or just a *component.* Graphs are usually represented by diagrams, points in the diagram corresponding to points of the graph, and a line segment joining points u and v in the diagram iff (u,v) is in E(G).

Thus for example the graph G whose diagram is shown in figure I.2.1. below

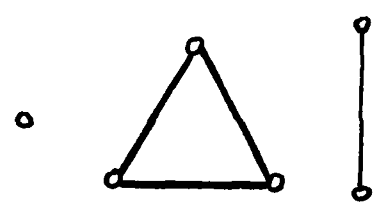


Figure I.2.1.

is disconnected; its components are simply the three connected "pieces": the isolated point, the triangle, and the line.

A circuit C_n with n points ($n \geq 3$) is the graph represented by an n-sided polygon.

We now define a special kind of graph, following the terminology in (2). A graph L is called a *linear graph* if each of its components is either a single line or a circuit. The graph G of I.2.1. is not a "linear graph" since one of its components is a point v. However, G - v, the graph obtained from G by deleting the point v and all lines adjacent to v, is a linear graph. If we wish to delete only one line x, we denote the resulting graph by G - x.

G is called *bipartite* if it is possible to partition V(G) into two nonempty classes such that no two vertices in the same class are adjacent. According to König's Theorem (11), G is bipartite iff no odd circuit is a subgraph of G.

A tree is a connected graph with no cycles. The number of edges in a tree T is |T| - 1 (3, Theorem 4.1).

II.1. Let G be a graph with p points $v_1,\ldots\ldots v_p$. The adjacency matrix $A(G) = (a_{ij})$ is defined to be the p x p matrix such that $a_{ij} = 1$ if (v_i,v_j) is an edge of G, and 0 otherwise. In particular, A(G) has zeroes along its main diagonal, since we have not allowed loops.

If a different ordering of the points of G is used, the resulting matrix is equal to PAP^{-1} for some permutation matrix P and so the same characteristic polynomial is obtained, since similar matrices have the same characteristic polynomial.

Collatz and Sinogowitz (1) used $\det(A - xI)$, the characteristic polynomial of A(G), in their paper. For convenience (as we shall see), we use $\det(A + xI)$. Henceforth this expression is what will be meant by the (characteristic) polynomial of a graph G, and we shall denote it by P(G,x).

II.2. *Elementary properties of the polynomial of a graph.*

II.2.1. Let G and H be vertex-disjoint (i.e. $V(G)\cap V(H)$ empty), and let E be their union[2]. Then

$$P(E,x) = P(G,x)P(H,x).$$

Proof. With a suitable labelling of the vertices, the matrix $A(E)+xI$ is easily seen to be the direct sum of the matrices $A(G)+xI$ and $A(H)+xI$. Hence

$$\begin{aligned} P(E,x) &= \det(A(E)+xI) \\ &= \det(A(G)+xI) \cdot \det(A(H)+xI) \\ &= P(G,x) \cdot P(H,x) \end{aligned}$$

---

(2) The *union* of two graphs A and B, denoted $A\cup B$, is the graph whose vertex set is $V(A) \cup V(B)$ and whose edge set is $E(A) \cup E(B)$.

II.2.2. If the connected components of G are $G_1 \ldots G_n$, $P(G,x) = \prod_1^n P(G_i,x)$.

Proof. Use induction and II.2.1.

Proposition 1.

II.2.3. Let G_1 and G_2 be vertex-disjoint, and form H by adding to the union of G_1 and G_2 an edge from a vertex v_1 in G_1 to a vertex v_2 in G_2. Then

$$P(H,x) = P(G_1,x)P(G_2,x) - P(G_1 - v_1,x)P(G_2 - v_2,x).$$

Proof. Without loss of generality, we construct the matrix $A(H) + xI$ as follows: the block consisting of the intersection of the first $|G_1|$ rows and the first $|G_1|$ columns is precisely $A(G_1) + xI$, where the row corresponding to v_1 is the $|G_1|^{th}$ row in $A(H) + xI$. Then the block consisting of the last $|G_2|$ rows intersected with the last $|G_2|$ columns is precisely $A(G_2) + xI$, where the row corresponding to v_2 is the $(|G_1| + 1)^{th}$ row in $A(H) + xI$. In addition, the fact that v_1 and v_2 are joined results in a 1 being in the $(|G_1|, |G_1| + 1)$ and $(|G_1| + 1, |G_1|)$ positions. Elsewhere the entries are zero (see figure II.3.4.).

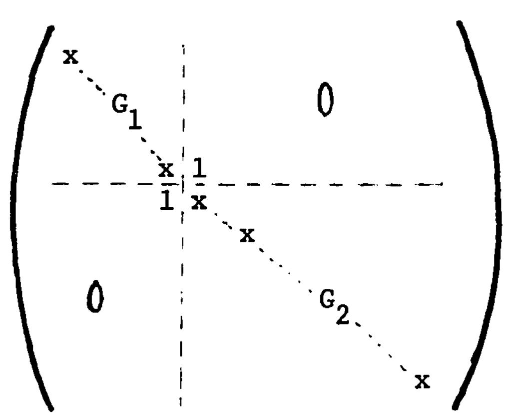


Figure II.3.4.

Let us evaluate the determinant of this matrix by the Laplace Expansion (12, p.14) using the first $|G_1|$ rows. Put $|G_1| + |G_2| = k$.

Let D be the set of all subsets of $\{1,2,\ldots k\}$ with $|G_1|$ elements. Let $d = \{1,2,\ldots |G_1|\}$, and for any $h \in D$, $h' = \{1,2,\ldots k\} - h$. Denote by $A_{f,g}$ the determinant of the matrix obtained from $A(H) + xI$ by retaining only the rows numbered in f and the columns numbered in g, where $f,g \subset \{1,2,\ldots k\}$ and $\mathrm{card}(f) = \mathrm{card}(g)$.

Define $\rho_{f,g}$ to equal $(-1)^v$ where there are v inversions between f and g; i.e. pairs (i,j) such that $i \in f$, $j \in g$, and $i > j$. Then $\det(A(H) + xI) = P(H,x) = \rho_{d,d'} \sum_{h \in D} \rho_{h,h'} \; A_{d,h} \cdot A_{d',h'}$ (Laplace).

Let us evaluate this sum. For a given term to be nonzero, h must take its $|G_1|$ columns from the first $|G_1| + 1$, otherwise $A_{d,h} = 0$. However, if h omits one of the first $|G_1| - 1$, $A_{d',h'} = 0$. Therefore there are only two choices of h: $\{1,2,\ldots |G_1|\}$ and $\{1,2,\ldots,|G_1| - 1, |G_1| + 1\}$.

Let us denote these h_1 and h_2. Now, $A_{d,h_1} = \det(A(G_1) + xI) = P(G_1,x)$. Similarly, $A_{d',h_1'} = P(G_2,x)$, $A_{d,h_2} = P(G_1-v_1,x)$ and $A_{d',h_2'} = P(G_2 - v_2,x)$. Also $\rho_{d,d'} = 1, \rho_{h_1,h_1'} = 1$, and $\rho_{h_2,h_2'} = -1$. Therefore $\det(A(H) + xI) = \rho_{d,d'} \Sigma \rho_{h,h'} \; A_{d,h} \cdot A_{d',h'}$ (only non-vanishing terms for $h = h_1$ or h_2) $= P(G_1,x)\, P(G_2,x) - P(G_1-v_1,x)\, P(G_2-v_2,x)$.

Q.E.D.

Using Theorem I, we shall give an easy combinatorial proof of this result in II.7.

II.3. Most published work concerning the adjacency matrices of graphs has involved the magnitude of, and bounds for, the least and greatest (the latter called the index) eigenvalues, and changes in these quantities under imbeddings[3]. We shall not be discussing these considerations.

Hoffman (8) defines "the polynomial of a graph G" to be a polynomial P of minimal degree that $P(A(G)) = J$, where J is the matrix of appropriate size consisting entirely of 1's. However, this is not related to the polynomial we treat here; in fact, Hoffman proves his polynomial exists iff G is regular and connected.

Harary (2) has conjectured (briefly) that two graphs G_1 and G_2 are isomorphic if their adjacency matrices A_1 and A_2 have the same set of eigenvalues (spectrum). However, as he states, several counterexamples have been found with graphs of sixteen points. He then goes on to ask what is the minimum number of points in any counterexample, and guesses sixteen. However, table II in the Appendix to (1) contains two different trees of eight points with identical spectra. These are shown in our Appendix.

---

(3) Hoffman's paper (7) has a comprehensive bibliography for this type of work.

II.4. *Example*

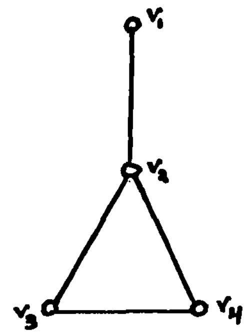


Figure II.4.1.

With the indicated labelling of the points of G (see Figure II.4.1.) we obtain the adjacency matrix A(G):

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix}$$

The polynomial of G, det(A + xI) is computed to be $x^4 - 4x^2 + 2x + 1$.
G has one spanning linear graph, shown in Figure II.4.2 .

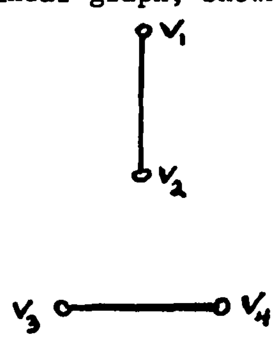


Figure II.4.2.

II.5. Given the polynomial of a graph, it is natural to ask what information can be deduced from the coefficients. Suppose the polynomial of a graph G is $P(G,x) = \sum_{i=0}^{p} a_i x^{p-i}$ where G has p points. Collatz and Sinogowitz (1) found the following geometric interpretations: (we refer to circuits of length 3, 4, and 5 as triangles, quadrilaterals, and pentagons respectively).

$a_0 = 1$

$a_1 = 0$ (the number of loops in G)

$-a_2 = q$, the number of edges in G.

$\frac{1}{2}a_3$ = the number of triangles in G (each set of 3 mutually-joined points is a triangle, and is counted *once*.)

a_4 = (the number of pairs of non-incident edges in G) - (twice the number of quadilaterals in G).

$-\frac{1}{2}a_5$ = (the number of pairs consisting of one triangle and a non-incident edge) - (The number of pentagons in G.)

We shall show in II.6. exactly how all coefficients arise, making use of some results by Harary (2), which we now summarize. Harary defines the *variable adjacency matrix* $A(G,Y) = (a_{ij})$ of a graph by assigning to each edge a variable y_k, and letting $a_{ij} = 0$ if v_1 and v_2 are not adjacent, and putting $a_{ij} = y_k$ if v_1 and v_2 are joined by a line, that line being y_k.

Here $Y = (y_1, y_2 \ldots)$.

The *variable determinant* of a graph is the determinant of its variable adjacency matrix. For example, the variable adjency matrix of the graph shown in Figure II.4.1, and again in Figure II.5.1 with its lines labelled is

$$\begin{pmatrix} 0 & y_4 & 0 & 0 \\ y_4 & 0 & y_1 & y_2 \\ 0 & y_1 & 0 & y_3 \\ 0 & y_2 & y_3 & 0 \end{pmatrix}$$

and its variable determinant is $y_3^2 \cdot y_4^2$.

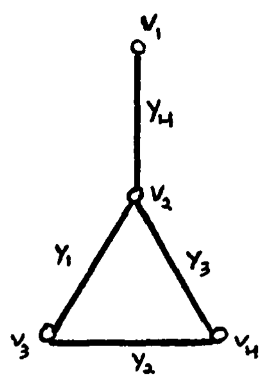


Figure II.5.1.

Harary proves that if the *spanning linear subgraphs* of G are $G_1 \ldots\ldots G_n$, then

II.5.2. $\det(A(G,Y)) = \Sigma \det(A(G_i,Y))$

When G does not have spanning linear subgraphs, det(A(G,Y)) is the empty sum, 0. Further, he proves:

II.5.3. $$\det(A(G_i,Y) = (-1)^{e_i} 2^{c_i} \prod_{y_k \in L_i} y_k^2 \prod_{y_j \in M_i} y_j$$

where e_i = number of even components in G_i

c_i = number of components in G_i which are circuits (more than two vertices)

L_i = set of components in G which are lines

M_i = set of remaining components of G_i (circuits)

For example, the graph of Figure II.5.1 has one spanning linear subgraph, shown in Figure II.5.4.

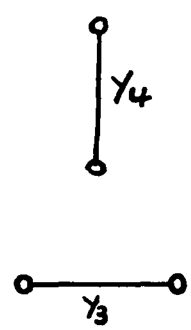


Figure II.5.4.

Applying II.5.2 and II.5.3 we deduce $\det(A(G,y)) = (-1)^2\, 2^0\, y_3^2\, y_4^2$, which is what we had calculated from the variable adjacency matrix.

Setting each $y_k = 1$ gives us det(A(G)), the constant term of the characteristic polynomial, i.e. P(G,0). In this case det(A(G)) = 1, as we had computed in II.4.

It is to be noted here that the only information about P(G,x) we can garner so far concerns the constant term. However, we shall extend these results in the next section so that the graph polynomial is completely determined by its linear subgraphs, and obtain Harary's result as a corollary.

II.6. We have just summarized the results in Harary's paper (2). He mentions, as we do, that the graphs he considers have no loops. He goes on to say that the extension to graphs having loops is straightforward; nowhere in his proofs is used the hypothesis that loops are not allowed, i.e., that the main diagonal of A(G) consists of zeroes. The only modification required is to the definition of a linear graph. Whereas in graphs without loops the nonzero terms in det(A(G,Y)) correspond to disjoint lines and circuits (which is what prompted the definition of a linear graph), graphs with loops will provide terms corresponding to isolated loops as well.

Therefore we define an *extended linear graph* to be a graph whose components are either loops, lines or circuits. Although we have defined a graph so as not include loops, we will use this extension of Harary's result to graphs with loops, but only in the proof of *Theorem I*.

Figure II.6.1 gives an example of a graph G with a loop, and Figure II.6.2 shows the four spanning extended linear graphs of G.

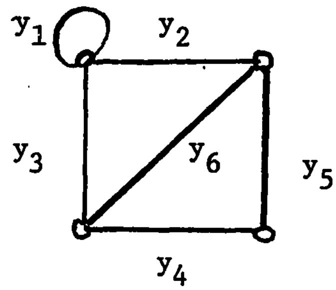


Figure II.6.1.

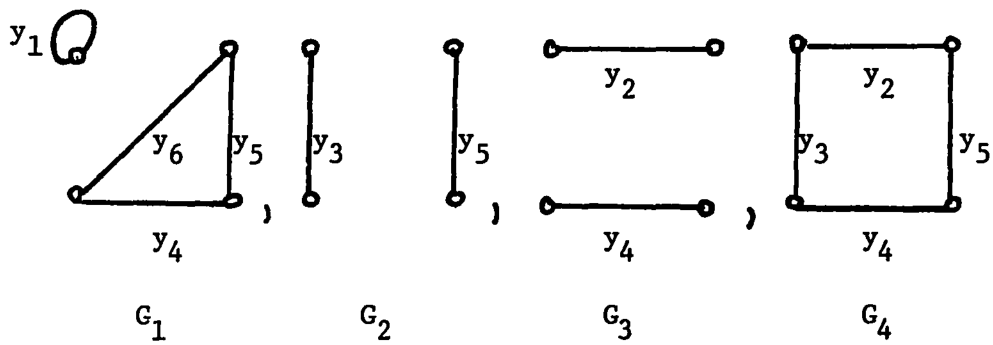


Figure II.6.2.

Using II.5.3 $\det(A(G_1,Y)) = +2(y_1)\ (y_4 \cdot y_5 \cdot y_6)$

$$\det(A(G_2,Y)) = +y_3^2 \cdot y_5^2$$

$$\det(A(G_3,Y)) = +y_2^2 \cdot y_4^2$$

$$\det(A(G_4,Y)) = -2y_2 \cdot y_3 \cdot y_4 \cdot y_5$$

We deduce from II.5.2 that

$$\det(A(G,Y)) = 2y_4y_5(y_1y_6 - y_2y_3) + y_3^2y_5^2 + y_2^2y_4^2$$

Setting each $y_i = 1$,

$$\det(A(G)) = 2$$

which can be verified directly from $A(G)$, which is

$$\begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix}$$

Definition. Let G be a graph.

Define $\sigma(G) = (-1)^e \, 2^c$ if G is an extended linear graph with e even components and c circuits (in particular $\sigma(\emptyset) = 1$), and $\sigma(G) = 0$ otherwise.

Theorem I. If G is a graph without loops, $P(G,x) = \Sigma\sigma(L)\, x^{|G|-|L|}$ *where the sum ranges over all the subgraphs L of G.*

Corollary (Harary). $P(G,0) = \Sigma\sigma(L)$ *where the sum ranges over all spanning linear subgraphs of G.*

Proof. As before, assign a variable y_k to each line of G. Furthermore, let us modify G by adding exactly one loop at each vertex. Call the new graph G^*, and assign the variable x_i to the new line forming the loop at the vertex v_i .

Then, applying Harary's extended result, we have

II.6.3. $$\det(A(G^*,Y)) = \sum_{1}^{s} (-1)^{e_i}\, 2^{c_i} \sum_{y_k \varepsilon L_i} y_k^2 \sum_{y_j \varepsilon M_i} y_j \sum_{x_n \varepsilon N_i} x_n$$

where G_1^*, G_2^*,.....G_s^* are the extended linear spanning subgraphs of G^*, N_i is the set of loops in G_i^*, and everything else is as previously defined in II.5.

Now let us set each $y_k = 1$ and each $x_n = x$. Then $\det(A(G^*,Y)) = P(G,x)$. It is then clear that what we now have from II.6.3 is

II.6.4 $$P(G,x) = \sum_1^s (-1)^{e_i} 2^{c_i} x^{n_i}$$

where n_i is the number of loops in G_i^*

Now, if a particular G_i^* has n_i loops, we can see that G_i^* minus these loops is a linear subgraph L_i of G containing $|G| - n_i$ vertices; i.e., $n_i = |G| - |L_i|$. Conversely, any linear subgraph L of G can be made into an extended spanning linear subgraph G_i^* of G^* by adding to L the loop corresponding to each vertex not contained in L (if indeed there are any to be added).

This estabishes a one-to-one correspondence between the L_i and the G_i^* (where $L_1, L_2 \ldots . L_s$ are *all* linear subgraphs of G). Note also that each corresponding pair L_i and G_i^* have the same number of even components and circuits, since a loop is an odd component and not a circuit.

$$\therefore \quad (-1)^{e_i} 2^{c_i} = \sigma(L_i)$$

Thus we can rewrite II.6.4 as

$$P(G,x) = \Sigma\sigma(L_i)\, x^{|G| - |L_i|}$$

where the sum ranges over the linear subgraphs of G, and we have used $n_i = |G| - |L_i|$. Since $\sigma(L) = 0$ unless L is a linear graph, we could just as well have the sum range over all subgraphs of G. The corollary is obtained by setting x equal to 0.

Q.E.D.

II.7. *Applications.*

Let us prove II.2.3 by means of Theorem I.

What are the linear graphs of H? One type consists of any linear graph from G_1, and any from G_2. (Letting $\emptyset$ be a linear graph). Note that if C is the disjoint union of A and B, $\sigma(C) = \sigma(A) \cdot \sigma(B)$. Now

$$P(G_1,x) \cdot P(G_2,x) = \left[\sum_{L_1 \varepsilon G_1} \sigma(L_1)\, x^{|G_1| - |L_1|}\right] \cdot \left[\sum_{L_2 \varepsilon G_2} \sigma(L_2)\, x^{|G_2| - |L_2|}\right]$$

$$= \sum_{L_1, L_2} \sigma(L_1 \cup L_2)\, x^{|H| - |L_1 \cup L_2|}$$

and therefore $P(G_1,x)P(G_2,x)$ is the contribution of this type of linear graph to $P(H,x)$.

The other type of linear graph in H has the line $(v_1,v_2) = x$ for one of its components. Then clearly any further components must be from $G_1 - v_1$ and $G_2 - v_2$ (and any will do).

The contribution of this type of linear graph to $P(H,x)$ is then $-P(G_1 - v_1,x)P(G_2 - v_2,x)$ where the minus sign appears because each such graph contains x, and $\sigma(x) = -1$.

$$\therefore \quad P(H,x) = P(G_1,x)P(G_2,x) - P(G_1 - v_1,x)P(G_2 - v_2,x)$$

Q.E.D.

In the case of a tree T, the only linear graphs contained in T, besides $\emptyset$, are sets of disjoint edges. This allows an easy interpretation of $P(T,x)$.

Let e_i be the number of different combinations of i disjoint edges in T. If L consists of j disjoint edges, $\sigma(L) = (-1)^j$. We therefore deduce from Theorem I

II.7.1 $$P(T,x) = \sum_{i=0} (-1)^i e_i x^{|T|-2i}$$

We shall be using this in section III. This shows a tree polynomial has either odd powers only, or even powers only. However, this characterizes not trees, but a larger class of graphs.

Theorem II. Let G be a connected graph. G is bipartite iff P(G,x) has even powers only or odd powers only (i.e. P(G,x) an even or odd function respectively).

Proof. Suppose G is bipartite. Then by König's Theorem (11, p.170) G has no odd circuits. Therefore G contains no linear subgraphs with an odd number of vertices; hence $P(G,x) = \Sigma\sigma(L)x^{|G| - |L|}$ (Theorem I) has powers only of the same parity as G.

Conversely, suppose the powers of x in P(G,x) are either all odd or all even. Then there is no term in $x^{|G|-3}$. Since this term has coefficient equal to twice the number of triangles (by Theorem I) there are no triangles in G.

Suppose there are no odd circuits of length $\leq j$ in G (j odd). Any linear subgraph of $j+2$ vertices must contain an odd circuit (since $j+2$ is odd). But there are no odd circuits of length $\leq j$. Hence the only possible linear subgraphs of $j+2$ vertices are circuits of length $j+2$, and the coefficient of $x^{|G|-j-2}$ is twice the number of such circuits. But the coefficient of $x^{|G|-j-2}$ is 0 by hypothesis. Hence there are no odd circuits of length $j+2$.

This induction shows that G has only even circuits, and hence is bipartite by the result of König cited in I.2.

Q.E.D.

III.1. One of the problems suggested by Collatz and Sinogowitz (1) was to find a geometric interpretation of graphs having 0 in their spectrum. He called such graphs "non-primitive". However, since tradition has not yet cemented this definition, we will use the term *singular* to refer to a graph G having 0 in its spectrum, i.e., such that A(G) is singular, or equivalently $P(G,0) = 0$. In this section we shall characterize singular trees and give some sufficient geometrical conditions for graphs to be singular; first some geometrical remarks and some definitions.

We define a *chain* of length n to be a point for $n=1$, a line for $n=2$, and for $n \geq 3$ the graph obtained by deleting any edge from C_n.

For $n \geq 3$, the two points originally joined by the deleted edge are called endpoints; for $n = 1$ or 2 all points are endpoints. Given a connected graph G, we say a chain of length n *stems* from v if there exists an edge (u,v) such that G-(u,v) has two components, one being a chain of length n with u as one of its endpoints. For instance, in Figure III.1.1 chains of lengths 2 and 3 stem from v.

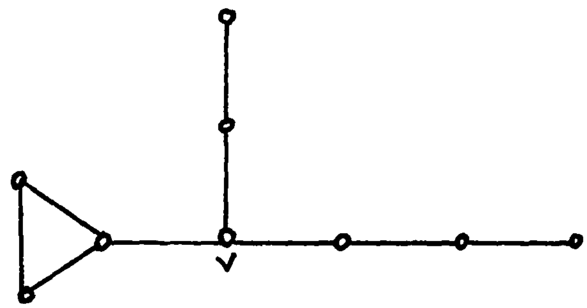


Figure III.1.1.

Given a vertex v in a tree T with k edges incident to it, it is easy to see that the fact that a tree has no circuits means that the remaining vertices of T are partitioned into k disjoint classes. (These are also said to *stem* from v.)

Definition: We say v is of type i $(i \geq 0)$ if i odd classes stem from v. The type of v is denoted t(v). Note that if T is odd, P(T,x) has only odd powers of x (Theorem II) and hence $P(T,0) = 0$ necessarily. Thus any odd tree is singular.

Lemma 1. The maximum cardinality of a set of disjoint edges in a tree T is $[\![|T|/2]\!]$ (4).

Proof. Suppose there exists a set with $[\![|T|/2]\!]+k$ edges. Then $|T| \geq 2\,([\![|T|/2]\!]+k) \rightarrow 2k \leq |T|-2[\![|T|/2]\!] \leq 1$

$$\rightarrow k \leq 0$$

Q.E.D.

Lemma 2. In any tree T, $|T| \geq 3$, there exists v such that at least two chains stem from v.

Proof. The case $|T| = 3$ is trivial. Assume the result for $|T| = n$ $(n \geq 3)$. Suppose $|T| = n+1$. Any tree has an endpoint (a point incident with only one edge) (3, Corollary 4.1A). Remove from T one endpoint e and the edge incident to it, obtaining T–e. Since $|T - e| = n$, T–e has a point v with the required property. Replace e and the edge incident to it. If v still has two or more chains stemming from it, we are through. Otherwise, e has been joined to a point p on a chain stemming from v and now p clearly has two chains stemming from it.

Q.E.D.

---

(4) $[\![x]\!]$ is the greatest integer less than or equal to x.

Theorem III. Let $|T|$ be even. The following are equivalent:

(i) T is non-singular

(ii) T has a set of $|T|/2$ disjoint edges

(iii) For every $v \in T$, $t(v) = 1$

Proof. We prove ii→iii, iii→ii and i↔ii:

ii→iii Clearly for any v, there must be at least one odd class stemming from v, since $|T|-1$ is odd (i.e., $t(v) \geq 1$). Suppose there are more than one for some v_o, so that $t(v_o) > 1$.

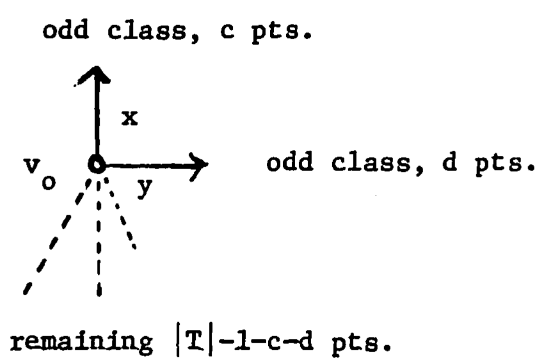


Figure II

See Figure II, where edges x and y both lead to sets of points of cardinalities c and d respectively, where c and d are supposed odd.

How large a set of disjoint edges can T contain? According to Lemma 1, the x-class can yield no more than $(c-1)/2$ disjoint edges, and the y-class no more than $(d-1)/2$. Also the remaining classes can contribute no more than $[(|T|-c-d-i)-1]/2$ edges. In addition, it is conceivable that one of the edges incident to v_o can be chosen. The maximum number of disjoint edges is then

$$(c-1)/2 + (d-1)/2 + (|T|-c-d-2)/2 + 1 = |T|/2 - 1$$

contrary to assumption.

<u>iii→ii</u> Induction on $|T|$ (the case $|T| = 2$ is trivial).

Assume iii→ii whenever $|T| = n$ (n even)

Now let $|T| = n + 2$

By Lemma 2, there exists a v_o from which at least two chains stem. Since $t(v_o) < 2$ by assumption, there is at least one even chain stemming from v_o. Choose one of these even chains, and delete from it the endpoint not adjacent to v_o and the point adjacent to this endpoint (and of course the two edges incident to the two deleted points). The resulting graph T' has n vertices, $t(v) = 1$ for all $v \varepsilon T$, and so a set of $n/2$ disjoint edges exists in T'. To this same set in T, add the previously deleted edge incident to the endpoint (this edge cannot be incident with T'). We now have a set of $(n + 2)/2$ disjoint edges.

Q.E.D.

ii↔i According to our interpretation of P(T,x) (II.7.1), P(T,x) has a nonzero constant term (i.e., P(T,0)≠0) iff there exists a set of $|T|/2$ disjoint edges.

Q.E.D.

Theorem III is related to a theorem of Tutte (14) concerning 1-*factors* (A set of $|G|/2$ independent lines in G is a 1-factor of G.)

Tutte's theorem states:

A graph G has a 1-factor iff $|G|$ is even and there is no set of points S such that the # of odd components of G–S exceeds card (S).

Applying this to an even tree T, we deduce that T can have a 1-factor (i.e., $|T|/2$ independent lines) only if for each $v \in T$, T–v has one odd component. This would then be an alternate way of proving the necessity of condition (iii) in Theorem III.

III.2 *Theorem IV. Suppose for every vertex v in T, t(v) = 1. Then* $P(T,0) = (-1)^{|T|/2}$

Proof. Let e be an endpoint of T. Since only one class stems from e, and t(e) = 1 by hypothesis, $|T|-1$ is odd and consequently $|T|$ is even.

By Theorem III, there exists at least one set of $|T|/2$ independent edges. Our interpretation of a tree polynomial (II.7.1) according to Theorem I tells us that $P(T,0) = (-1)^{|T|/2} e_{|T|/2}$ where $e_{|T|/2}$ is the number of different sets of $|T|/2$ independent edges.

Therefore there remains only to prove that $e_{|T|/2} \leq 1$, which we do by induction (for $|T|$ even).

The case $|T| = 2$ is trivial.

Assume for any tree T such that $|T| = 2n$ $(n \geq 1)$, we have $e_{|T|/2} \leq 1$. Now let $|T| = 2(n + 1)$.

Suppose T has two sets S_1 and S_2 of $|T|/2 = n + 1$ independent edges. We must show $S_1 = S_2$.

Let e be an endpoint of T, where x is the edge incident to e, and e' the vertex adjacent to e. If $x \notin S_1$, T-e has n + 1 independent edges. But Lemma I asserts T-e can have no more than n independent edges. Therefore $x \in S_1$. Similarly $x \in S_2$.

Let T' = T-e-e'. Since $|T'| = 2n$, the induction hypothesis asserts that T' can have no more than one set of n independent edges. Since none of the edges incident to e' except x can be in S_1 or S_2, $S_1 - \{x\}$ and $S_2 - \{x\}$ are both sets of n independent edges from T'. Therefore $S_1 - \{x\} = S_2 - \{x\}$ and it follows that $S_1 = S_2$.

Q.E.D.

Theorems III and IV settle Collatz' proposed problem in the case of trees, but we have found no such characterization for singularity of general graphs. It seems unlikely that one exists, since $P(G,0) = (-1)^{e_i} 2^{c_i}$ and this sum happens to "cancel out" apparently at random. Perhaps further progress in this direction can only hope to proceed on special kinds of graphs (as we did on trees). However, there are sufficient but not necessary conditions under which we can state that a graph is singular.

Proposition 2. Suppose G satisfies one of these conditions:

(a) There exists a vertex v_o from which stem at least two odd chains, or

(b) There exist two unjoined vertices v_1 and v_2 which are adjacent to exactly the same vertices.

Then G is singular.

Proof. (a) We will show such a graph can have no spanning linear subgraphs, and hence $P(G,0) = 0$

Let two odd chains stemming from v_o have as vertex sets $\{v_1,\ldots\ldots v_i\}$ and $\{v_1'\ldots\ldots v_j'\}$ (i, j odd) where v_1 and v_1' are adjacent to v_o, v_k adjacent to v_{k-1} and v_{k+1} $(2\leq k\leq i-1)$ and v_k' adjacent to v_{k-1}' and v_{k+1}' $(2\leq k\leq j-1)$

If there is to be a spanning linear subgraph L, it must contain the vertex v_i, and it can only do so if the edge (v_{i-1}, v_i) is a component of L. Similarly v_{i-2} can be in L only if the edge (v_{i-3},v_{i-2}) is a component of L. Eventually we reach the conclusion that (v_o,v_1) must be a component of L. Similarly, (v_o,v_1') must be a component of L.

But this is impossible since these two edges are not disjoint.

(b) In the matrix A(G), the two rows (or columns) corresponding to v_1 and v_2 are the same, hence $\det(A(G)) = P(G,0) = 0$.

Q.E.D.

III.3 A point and a line are said to *cover* each other if they are incident.

A set of points which covers all the lines of a graph G is called a *point cover* for G, while a set of lines which covers all the points of G is a *line cover*.

The smallest number of points in any point cover for G is called its *point covering number* and is denoted $\alpha_o(G)$ or α_o . Similarly $\alpha_1(G)$ or α_1 is the smallest number of lines in any line cover of G and is called its *line covering number*.

The largest number of mutually non-adjacent points in G is called the *point independence number* of G, denoted $\beta_o(G)$ or β_o. The largest number of independent (vertex-disjoint) lines in G is the *line independence number* $\beta_1(G)$ or β_1.

Gallai (see 3, Theorem 10.1) proved:

III.3.1. For any nontrivial connected graph G,

$$\alpha_o + \beta_o = \alpha_1 + \beta_1 = |G|$$

König (10) proved:

III.3.2. If G is bipartite, $\beta_1 = \alpha_o$

We shall now see that P(G,x) can in certain cases yield information regarding these numbers.

Theorem V. Let the lowest power of x to appear in P(G,x) be x^d.

(a) If G is a tree, $\alpha_1 = \beta_o = \frac{1}{2}(|G| + d)$

$$\alpha_o = \beta_1 = \tfrac{1}{2}(|G| - d)$$

(b) If G is bipartite, $\alpha_1 = \beta_o \leq \frac{1}{2}(|G| + d)$

$$\alpha_o = \beta_1 \geq \tfrac{1}{2}(|G| - d)$$

Proof. (a) From our previous interpretation of a tree polynomial (II.7.1), the last term in $P(G,x)$ is $(-1)^i e_i x^{|G|-2i}$ where i is the largest number of independent lines in G. By definition, $i = \beta_1(G)$, and by hypothesis $d = |G|-2i$. We deduce $\beta_1(G) = \frac{1}{2}(|G|-d)$. The other equations follow from III.3.1 and III.3.2.

(b) Since there is a term cx^d in $P(G,x)$ we know by Theorem I that there must be at least one linear subgraph L of G with $|G|-d$ vertices. By Theorem II, we know L consists of lines and/or even circuits. From any even circuit with k vertices, it is possible to extract k/2 independent lines. Hence from L we can derive a set of $\frac{1}{2}(|G|-d)$ independent lines. Hence $\beta_1(G) \geqslant \frac{1}{2}(|G|-d)$. Once again the other equations follow from III.3.1 and III.3.2.

Q.E.D.

The problem of finding a maximal set of independent lines in a graph(5) has been the subject of much investigation(6) .

---

(5) This is usually called the "maximum matching" problem.

(6) See for instance Chapter 7 in *Theory of Graphs* by O. Ore, Amer. Math. Society, Providence, 1962.

Although algorithms for obtaining such sets have been developed[7], no formulas for the number of lines in such sets (i.e., β_1) seem to have been published. Theorem V yields such a formula for trees, as well as a lower bound in the case of bipartite graphs. The basic data required is the adjacency matrix of the graph.

---

[7] e.g. M.L. Balinski, "*Labelling to obtain a maximum matching*", appears in *Combinatorial Mathematics and its applications*, Univ. of North Carolina Press, Chapel Hill, 1969.

IV.1. *Ulam's Conjecture*

For any graph G, there are $|G|$ graphs of the form G-v, one for each vertex v in G. Ulam's well-known conjecture (15) in its graph-theoretical form states that this collection of $|G|$ graphs uniquely determines G. Formally, let G have points $\{v_i\}$ and H have points $\{u_i\}$ with $|G| = |H| \geq 3$. If for each i the graphs $G_{(i)} = G-v_i$ and $H_{(i)} = H-u_i$ are isomorphic, then the graphs G and H are isomorphic. The graphs $G_{(i)}$ we call the *Ulam subgraphs* of G. Kelly (9) has succeeded in proving Ulam's conjecture for trees[8].

If we label the edges of G by $x_1.....x_q$, the *line form* of Ulam's conjecture states that G is characterized by the q graphs $G^{(i)} = G-x_i$.

A problem intimately related to Ulam's conjecture is that of *reconstruction*. Given n graphs $G_1, G_2,.....G_n$ of n-1 points each, when can we find a graph G (called a reconstruction) with n points $v_1.....v_n$, such that $G_i = G-v_i$ $(1 \leq i \leq n)$. Ulam's conjecture can then be stated: Given such a set of graphs there exists at most one reconstruction for it.

The current state of knowledge concerning reconstruction is summarized in (13).

([8]) Kelly's result antedated Ulam's conjecture and is generally thought to have motivated it. In (9) Kelly verified the conjecture for graphs with up to six points; and Harary and Palmer in (6) for graphs of seven points.

IV.2 *Definition:* The two graphs G and H are *Ulam-related* if for each i (with a suitable ordering) $G_{(i)} \simeq H_{(i)}$.

Note that the definition impies two Ulam-related graphs have the same number of vertices. If instead $G^{(i)} \simeq H^{(i)}$ for each i, we call the graphs G and H *Ulam-line-related.*

It is important to realize that if we are considering a set of (possibly non-distinct) graphs $\{G_i\}(1 \leq i \leq h)$ and searching for a reconstruction G, the graphs G_i are not jointly labelled. For instance, we have no way of determining, (in general) which vertices in G_1 are which in G_2 . It is uncertain whether or not graph polynomials can be of any use in proving (or disproving) Ulam's conjecture. However, we shall show how they can yield circumstantial evidence and how they indicate that graphs with certain properties might be proven to obey Ulam's conjecture. For instance, since Ulam's conjecture holds for trees, we should be able to prove, and we do,that two Ulam-related trees have the same polynomial. If we then find other types of graphs for which being Ulam-related implies having the same polynomial, these types of graphs seem like good candidates to satisfy Ulam's conjecture. Of course, they may not, since we have not proven that two graphs which are Ulam-related and have the same polynomial are isomorphic.

This last statement, if proven, would, as we shall see, yield several classes of graphs satisfying the conjecture. It may be worthy of further investigation.

We shall show that when G and H are Ulam-related, P(G,x) and P(H,x) are quite similar and possibly always the same. If a case were found where these polynomials were different, we would have a counterexample for Ulam's conjecture, since isomorphic graphs have identical polynomials (only the labelling is different).

In addition, we shall show that two Ulam-line-related graphs always have identical polynomials. This may indicate that the line-form of the conjecture is a simpler problem.

IV.3 We now prove that the polynomials of two Ulam-related graphs differ by a constant.

Theorem VI. Let the graph G have Ulam subgraphs $G_{(1)}, \ldots, G_{(n)}$. *Then for some constant c,*

IV.3.1 $$P(G,x) = \int_0^x \sum_1^n P(G_{(i)},t)\, dt + c$$

Proof. Consider a linear subgraph L of G with $|L| < |G|$. The graph L is a subgraph of a particular $G_{(i)} = G - v_i$ iff the vertex v_i is not contained in L. Thus, L is a subgraph of exactly $|G| - |L|$ Ulam subgraphs of G. If L has $|G|$ vertices, it is not a subgraph of any $G_{(i)}$. Let us then consider the expression

IV.3.2 $$\sum_{i=1}^{n} \sum_{L\varepsilon G_{(i)}} \sigma(L)x^{n-|L|}/(n-|L|)$$

where the second sum is over all linear subgraphs of $G_{(i)}$ including $\emptyset$ and $G_{(i)}$ itself.

From Theorem I,

$$P(G_{(i)},t) = \sum_{L\varepsilon G_{(i)}} \sigma(L)t^{|G_{(i)}|-|L|}$$

$$= \sum \sigma(L)t^{n-|L|-1}$$

We deduce $\int_0^x P(G_{(i)},t)dt = \sum \sigma(L)x^{n-|L|}/(n-|L|)$

Substituting in IV.3.2 we get the equivalent expression

IV.3.3 $$\sum_{1}^{n} \int_0^x P(G_{(i)},t)dt$$

Now let us return to IV.3.2.

Bearing in mind that a subgraph of G with $j < n$ vertices is a subgraph of exactly $n - j$ Ulam subgraphs, we see that IV.3.2 is equal to

IV.3.4 $$\sum_{\substack{L\varepsilon G \\ |L|<|G|}} \sigma(L)\ x^{|G|-|L|}$$

But this is precisely $P(G,x) - P(G,0)$ (Theorem I and its corollary).

We therefore have an equality between P(G,x) - P(G,0) and the expression IV.3.3:

Thus $$P(G,x) - P(G,0) = \Sigma \int_0^x P(G_{(i)},t)\,dt$$

or $$P(G,x) = \int (\Sigma P(G_{(i)},t))dt \quad + \quad c$$

where c = P(G,0) Q.E.D.

Corollary $$\frac{d}{dx} P(G,x) = \Sigma P(G_{(i)},x)$$

Theorem VII.

Let G and H be Ulam-related graphs satisfying any one of these conditions:

(a) Either G or H is known to be a tree of at least three points.

(b) G and H are both singular.

(c) det A(G) = det A(H).

(d) G and H each have a pair of adjacent vertices which are adjacent to exactly the same points.

(e) G and H each have a pair of non-adjacent vertices which are adjacent to exactly the same vertices

(f) G and H each have a vertex from which stem at least two odd chains.

Then $$P(G,x) = P(H,x)$$

Proof. By Proposition 2 (II.2) conditions (e) and (f) imply that $P(G,0) = P(H,0) = 0$. In addition, Theorem VI implies that $P(G,x)$ and $P(H,x)$ differ by a constant. We therefore deduce $P(G,x) = P(H,x)$. The same reasoning yields the sufficiency of conditions (b) and (c).

If G and H satisfy (d), the matrices corresponding to $P(G,1)$ and $P(H,1)$ each have two identical rows. Thus, $P(G,1) = P(H,1) = 0$, and again we conclude that $P(G,x) = P(H,x)$.

Proof of (a):

Suppose G is a tree. Then there are at least two connected Ulam subgraphs (corresponding to the removal of an endpoint). Therefore H must be connected, or else two Ulam subgraphs of H never could be connected (since $|H| = |G| \geqslant 3$). Also, H has the same number of edges as G (by Theorem VI). Hence H is also a tree. If $|G| = |H|$ is odd, we know from Theorem II that $P(G,0) = P(H,0) = 0$ and hence that $P(G,x) = P(H,x)$. If G is even, we examine the Ulam subgraphs $\{G_i\}$ (which are the same as the $\{H_i\}$). We conclude from Theorem IV that $P(G,0) = P(H,0) = 0$ if some G_i has more than one odd component, and that $P(G,0) = P(H,0) = (-1)^{|G|/2}$ otherwise. In either case it then follows from Theorem VI that $P(G,x) = P(H,x)$.

Q.E.D.

Theorem VII suggests six types of graphs for which it may be possible to prove Ulam's conjecture. As mentioned, Kelly (9) has proven it for trees. In the above proof for (a), we could have appealed to this result after proving H was necessarily a tree, deduced $G \simeq H$ and hence $P(G,x) = P(H,x)$.

Finally, we prove two Ulam line-related graphs have identical polynomials, or equivalently:

Proposition 3. Let G be a non-linear graph with q lines and with Ulam line-subgraphs $G^{(1)}, \ldots, G^{(q)}$. Then $P(G,x)$ is given by

IV.3.5 $$\sum_{1}^{q} \sum_{L \in G^{(i)}} \sigma(L) x^{|G|-|L|} / (q - e(L))$$

where $e(L) =$ the number of edges in L

Proof. Any subgraph L of G with $e(L)$ edges is a subgraph of exactly $q-e(L)$ Ulam line-subgraphs. Thus the sum in IV.3.5 is equal to $\sum_{L \underset{\neq}{\in} G} \sigma(L) x^{|G|-|L|}$.

By Theorem I, this equals $P(G,x) - \sigma(G)$.

By hypothesis however, $\sigma(G) = 0$. Hence we deduce IV.3.5.

Q.E.D.

IV.4. Theorem VI may furnish a useful tool in the problem of reconstruction defined in IV.1.

Given a collection of n graphs $G_1 \ldots . G_n$ with n-1 points each, an existence problem arises. Do these graphs admit a reconstruction? Very little progress has been made on this problem. If there exists a reconstruction G, we can easily determine what its polynomial Q(x) should be (up to a constant) using Theorem VI.

Thus if we had a set of necessary conditions for a polynomial to be a graph polynomial, we could apply this knowledge to see whether Q(x) can be a graph polynomial. This problem seems to be untouched, however.

Example. Do the five graphs of Figure IV.4.1 have a reconstruction G?

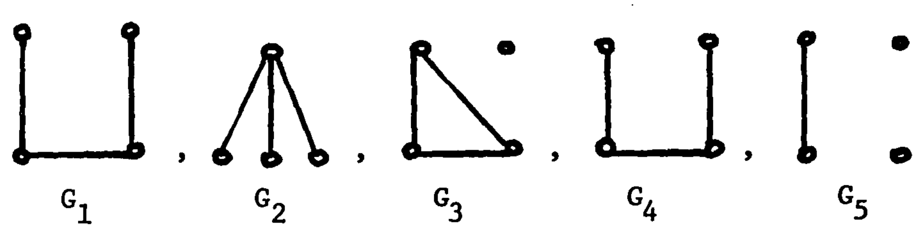


Figure IV.4.1.

Using Theorem I we calculate:

$$P(G_1,x) = P(G_4,x) = x^4 - 3x^2 + 1$$

$$P(G_2,x) = x^4 - 3x^2$$

$$P(G_3,x) = x^4 - 3x^2 + 2x$$

$$P(G_5,x) = x^4 - x^2$$

Thus, if G exists,

$$P(G,x) = \int_0^x (\Sigma P(G_i,t))dt + c$$

$$= \int_0^x (5t^4 - 13t^2 + 2t + 2)dt + c$$

$$= x^5 - 13/3\, x^3 + x^2 + 2x + c$$

But this is clearly not a graph polynomial, hence no reconstruction exists.

References

1. Collatz L., and Sinogowitz, U., *Spectren endlicher Grafen*, Abhandlung aus dem Mathematischen Seminar der Universität, Hamburg, 21 (1957),63-77.

2. Harary, F., *The determinant of the adjacency matrix of a graph*, SIAM Review 4 (1962),202-210.

3. __________ *Graph theory*, Addison Wesley, Reading, 1969.

4. __________ *Graphical Reconstruction*, Chapter V in *A seminar on graph theory*, (F. Harary, ed.), Holt, Rhinehart, New York, 1967.

5. __________ and Palmer, E.M., *The reconstruction of a tree from its maximal subtrees*, Canad. J. Math 18 (1966), 803-810.

6. __________ and Palmer, E.M., *On similar points of a graph*, J. Math. Mech. 15 (1966), 623-630.

7. Hoffman, A.J., *The eigenvalues of the adjacency matrix of a graph*, Chapter 32 of *Combinatorial Mathematics and its applications* (R.C. Bose, T.A. Dowling, ed.), Univ. of North Carolina Press, Chapel Hill, 1969.

8. __________ *On the polynomial of a graph*, Amer. Math Monthly 70 (1963), 30-36.

9. Kelly, P.J., *A congruence theorem for trees*, Pacific J. Math. 7 (1957), 961-968.

10. Konig, D., *Graphen und Matrizen*, Mat. Fiz. Lapok 38 (1931), 116-119.

11. __________ *Theorie der endlichen und unendlichen Graphen*, Leipzig (1936), Reprinted Chelsea, New York (1950).

12. Minc, H., and Marcus, M., *A survey of matrix theory and matrix inequalities*, Prindle, Weber and Schmidt, Boston, 1964.

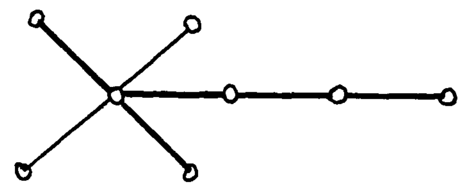
13. O'Neil, P.V., *Ulam's conjecture and graph reconstruction*, Amer. Math. Monthly 77 (1970), 35-43.

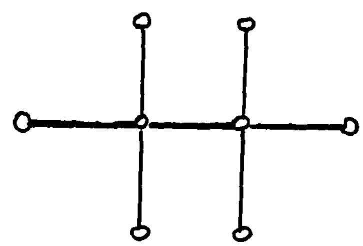
14. Tutte, W.T., *The factorization of linear graphs*, J. London Math. Soc. 22 (1947), 107-111.

15. Ulam, S.M., *A collection of mathematical problems*, Wiley, 1960.

A P P E N D I X

An Example of Two Graphs with Identical Spectra





We find that each of these graphs has:

7 edges

9 pairs of disjoint edges

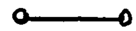
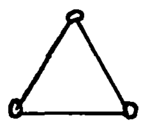
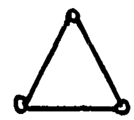
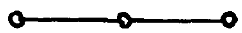
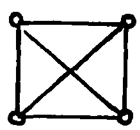
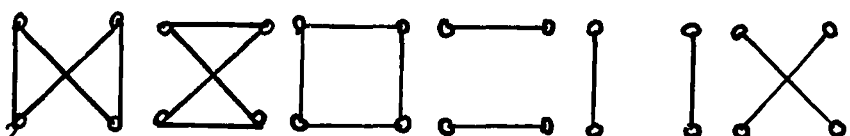
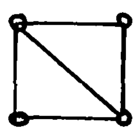
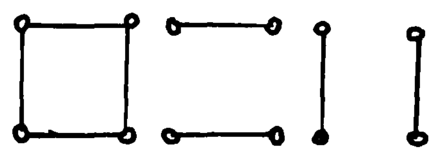
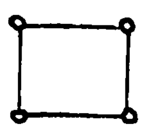
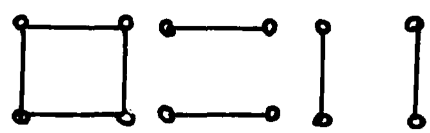
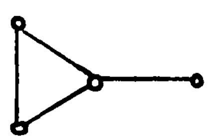
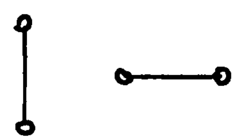
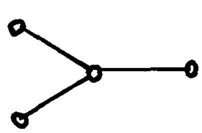
0 sets of $k \geq 3$ disjoint edges

Using II.7.1 the common polynomial is computed:

$$x^8 - 7x^6 + 9x^4$$

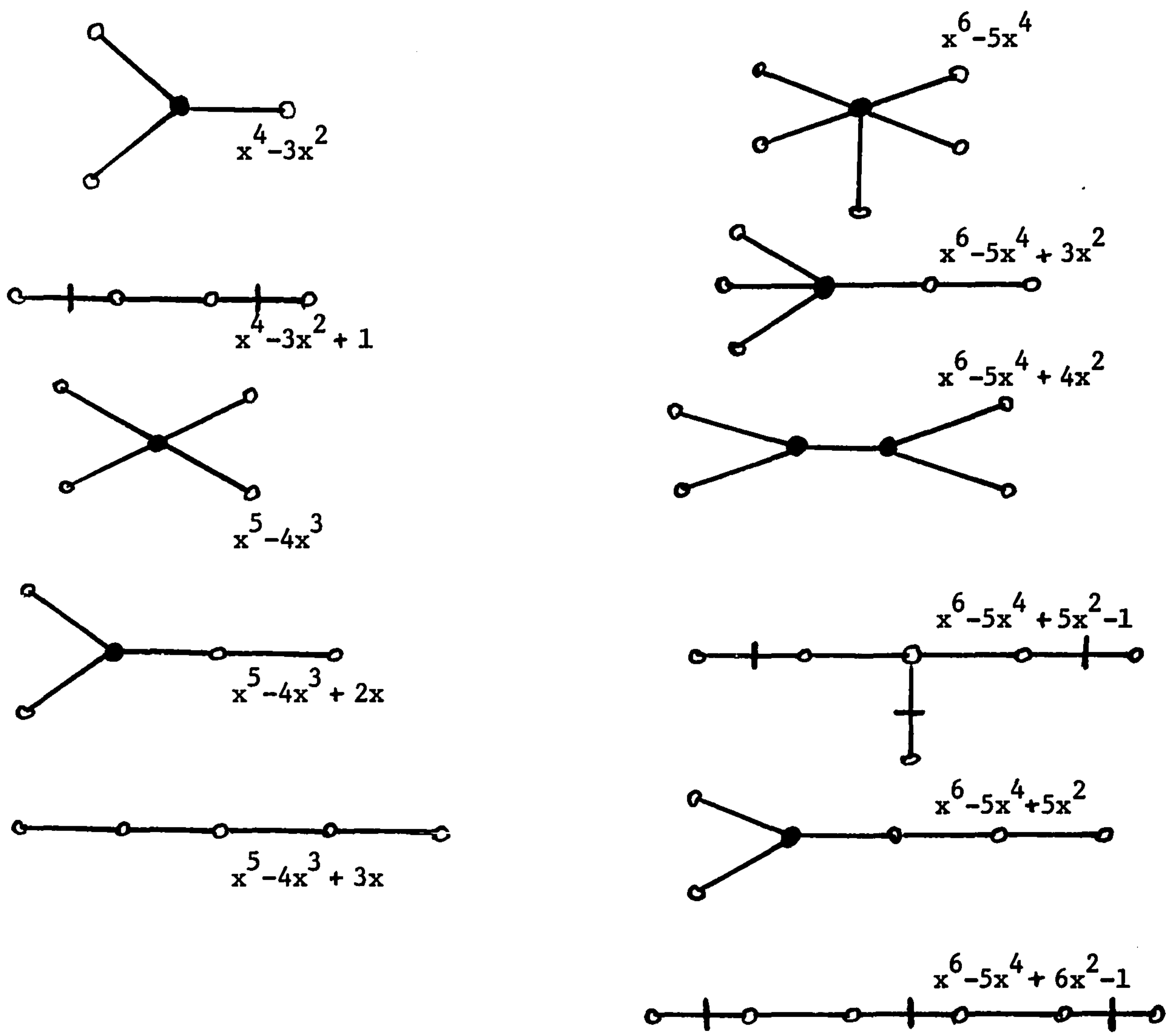
T A B L E I

The connected graphs with up to four points, with their polynomials and spanning linear subgraphs, if any.

	x	
	x^2-1	
	x^3-3x+2	
	x^3-2x	
	x^4-6x^2+8x+3	
	x^4-5x^2+4x	
	x^4-4x^2	
	x^4-4x^2+2x+1	
	x^4-3x^2	
	x^4-3x^2+1	

T A B L E II

The trees with 4, 5 and 6 points, and their polynomials. An even tree either has at least one vertex of type greater than 1 (shaded in the diagram) or a 1-factor. In the latter case the edges of the 1-factor are indicated.



Some further tabulations may be found in (1). Collatz and Sinogowitz have listed for each connected graph with up to five points, a polynomial[9] which in our notation is ± P(G,-x), and the roots of this polynomial. They also list this polynomial for the trees with 6, 7 and 8 vertices.

---

(9) See II.1