

THE DISSIMILARITY CHARACTERISTIC OF LINEAR GRAPHS

FRANK HARARY AND ROBERT Z. NORMAN¹

1. **Introduction.** The notion of dissimilarity characteristic was first introduced by Otter [5] in his theorem, "In any tree the number of nonsimilar vertices minus the number of nonsimilar lines (symmetry line excepted) is the number one." This concept was extended in [2] to a special kind of graph called a Husimi tree. The object of this note is to present a generalization to arbitrary graphs. A further extension to arbitrary simplicial complexes is being considered. We wish to acknowledge that our interest in these problems was inspired by Uhlenbeck [7], and we wish to thank him for suggestions which improved the presentation here.

Otter used his theorem above as a basic lemma for counting trees in terms of the number of rooted trees. Harary and Uhlenbeck [3] have utilized the dissimilarity characteristic for Husimi trees in order to obtain a functional equation for the number of Husimi trees in terms of the number of rooted ones, which they had already counted using the methods of Pólya [6]. Formulas are known [1] which give the number of graphs and of rooted graphs having any given number of points and lines. The equation developed here may be useful in finding a functional equation between these two generating functions. In addition it is hoped that this equation will prove useful in deriving the counting functions for the particular sorts of graphs discussed in [7].

2. **Definitions and motivation.** We begin by defining a multiply rooted graph, and shall then call it, for brevity, a graph. We assume that the definition of an ordinary or unrooted graph, as given in [4] for example, is known.

DEFINITION 1. A *multiply rooted graph* G is one in which the set of points of G is partitioned into disjoint distinguished classes Π_i . Two graphs G and G' are *isomorphic* if there exists a one-to-one correspondence between their point sets Π and Π' which preserves incidence and induces a one-to-one correspondence between their distinguished classes Π_i and Π'_j . An *automorphism* of G is an isomorphism of G

Presented to the Society, April 24, 1953; received by the editors May 18, 1953.

¹ The preparation of this paper was supported in part by a grant from the Rockefeller Foundation to the Research Center for Group Dynamics, and in part by a grant from the Office of Naval Research through the Engineering Research Institute of the University of Michigan.

onto itself which preserves distinguished classes. Let \mathfrak{R} denote the group of automorphisms of G . Two points of G are *similar* if an element of \mathfrak{R} sends one point into the other. Similarity of two lines and of two cycles are defined analogously.

DEFINITION 2. An *exceptional point of a cycle* (called symmetry point in [2]) is one incident to two similar lines of the cycle. An *exceptional line of a graph* (called symmetry line in [2]) is one whose end points are similar.

The dissimilarity characteristic of Husimi trees [2], a refinement of the Euler-Poincaré characteristic, was the starting point for the present investigation. By definition [7], a Husimi tree is a connected graph in which no line appears in more than one cycle. Let H be a Husimi tree; p , k , and c be the number of dissimilar points, lines, and cycles of H ; a be the number of exceptional lines of H not contained in any cycle; c_{pp} and c_{kk} be the number of dissimilar cycles with exactly two similarity classes of exceptional points and lines respectively. Then it was shown in [2] that the dissimilarity characteristic equation

$$(1) \quad p - (k - a) + (c - c_{pp} + c_{kk}) = 1$$

holds for any Husimi tree H .

Let c_{pk} be the number of dissimilar cycles containing exactly one class of exceptional points and one class of exceptional lines. We can reformulate (1) by introducing k_e as the number of dissimilar exceptional lines of H , and c_e as the number of dissimilar cycles having an exceptional point or an exceptional line. Clearly, $k_e = a + c_{pk} + 2c_{kk}$ and $c_e = c_{pp} + c_{pk} + c_{kk}$. From these last two relations and (1) we obtain

$$(2) \quad p - (k - k_e) + (c - c_e) = 1$$

as an alternative expression for the dissimilarity characteristic of Husimi trees. We shall show that with appropriate generalized definitions of the symbols appearing in (2), this equation holds for all connected graphs.

For the purpose of defining c and c_e , we shall consider chains on the oriented lines of G with integral coefficients. Then we define linear dependence with respect to similarity of a chain on a set of cycles of G . A dissimilarity cycle basis B is defined in terms of this dependence. Special consideration is given to those cycles of G having similar lines occurring with both positive and negative coefficients, and a certain well-defined subset of these will be called exceptional cycles. The numbers c and c_e will be defined as the number of cycles and exceptional cycles of B respectively. The precise definitions follow.

For any point v and any line λ of G , let $[v]$ and $[\lambda]$ denote the respective similarity classes. For each line of G , choose an arbitrary orientation subject to the restriction that similar lines have the same orientation with respect to similarity (with exceptional lines oriented arbitrarily). Further, let each cycle of G have an arbitrary orientation with similar cycles oriented similarly.

Let L be a chain on the oriented lines of G with integral coefficients. Let $L[\lambda]$ and $\bar{L}[\lambda]$ denote the sum of the positive and negative coefficients respectively of the lines of $[\lambda]$ in L .

DEFINITION 3. A chain L is *linearly dependent with respect to similarity* on a set of cycles z_1, z_2, \dots, z_m of G (written L is *dep* z_1, z_2, \dots, z_m) if there exist integers $a_0, a_1, a_2, \dots, a_m$ not all zero, such that for each similarity class $[\lambda]$ of lines of G ,

$$a_0 L[\lambda] = \sum_{i=1}^m a_i z_i[\lambda] \quad \text{and} \quad a_0 \bar{L}[\lambda] = \sum_{i=1}^m a_i \bar{z}_i[\lambda].$$

DEFINITION 4. A *dissimilarity cycle basis* B of G is a minimal collection of cycles such that every cycle of G is *dep* B . It is easy to show that given any cycle basis C of G there exists a dissimilarity cycle basis B which can be extracted from C . Equivalently, we may define a dissimilarity cycle basis B as a maximal collection of cycles independent with respect to similarity such that every cycle of G is *dep* B .

DEFINITION 5. A cycle z is *normal* if it contains no exceptional lines nor pairs of similar lines having opposite orientations with respect to the cycle. Thus z is normal if it is orientable with respect to similarity. Let N be the set of all normal cycles of G . A cycle z is *regular* if it is *dep* N ; otherwise z is *exceptional*.

Given a dissimilarity cycle basis B of G , let c and c_e denote the number of cycles and exceptional cycles of B respectively. It can readily be shown that c is an invariant of G . In §3 it will be shown that $c - c_e$ is invariant; thus c_e is also.

3. Proof of the theorem.

THEOREM. *Let G be a connected graph; B any dissimilarity cycle basis of G ; c the number of cycles of B ; c_e the number of exceptional cycles of B . Then equation (2), $p - (k - k_e) + (c - c_e) = 1$, holds for G .*

The idea of the proof is to transform the given graph G into a graph G' in which all points and lines are dissimilar. This is accomplished by first subdividing each exceptional line of G , thereby obtaining the intermediate graph G^0 . All points so added will be in new similarity classes distinct from those of G , and two of them will be similar if

they subdivide similar lines. Let G' be the graph obtained from G^0 by the identification of similar points and of similar lines, i.e., G' is the graph whose points and lines are classes of similar points and lines of G^0 . Let T denote the mapping² sending G onto $G' = T(G)$, where T sends points and exceptional lines of G onto points of G' , and sends all lines of G onto lines of G' . We note the dual role played by exceptional lines. It can then be shown that the Euler-Poincaré characteristic applied to the transformed graph G' yields the dissimilarity characteristic equation for G .

LEMMA 1. *If z' is a cycle of G' , then there exists a cycle z of G such that $T(z) = kz'$ for some positive integer k .*

PROOF. Let $[v_1], \dots, [v_t]$ be successive points of z' . Let u_1 be a point of G^0 in $[v_1]$. There exists a point u_2 in $[v_2]$ such that (u_1, u_2) is a line of G^0 , \dots , a point u_{t+1} in $[v_1]$ such that (u_t, u_{t+1}) is a line of G^0 , etc. Since the number of points in any similarity class of a finite graph is finite, there exists a minimal positive integer k such that $u_i = u_{kt+i}$. Then the set of points $u_i, u_{i+1}, \dots, u_{kt+i-1}, u_{kt+i}$ and the lines successively joining them form a cycle of G^0 , whose corresponding cycle z of G has the property $T(z) = kz'$.

Let B' be a cycle basis of G' . To each cycle z'_i of B' , associate any cycle z_i of G such that there exists an integer $k > 0$ for which $T(z_i) = kz'_i$ in accordance with Lemma 1. Let B^* be a set of cycles z_i of G thus associated with B' . The fact that B^* and B' have the same number of cycles will be used later.

LEMMA 2. *A cycle z of G is $dep B^*$ if and only if it is regular.*

PROOF. Once it has been shown that every cycle of N is $dep B^*$, the lemma follows at once. For since every regular cycle is by definition $dep N$, it would then be $dep B^*$. And since clearly B^* is a subset of N , every cycle $dep B^*$ is regular, proving the lemma.

We now show that every cycle of N is $dep B^*$. Let z be an arbitrary cycle of N , and let t denote its length. Let v_1 be any point of z , λ_1 the line (v_1, v_2) of z . Let λ_2 be the line (v_2, v_3) of z , $v_3 \neq v_1$. For each i , $2 \leq i \leq t-1$, define λ_i to be the line (v_i, v_{i+1}) , where $v_{i+1} \neq v_{i-1}$. Let λ_t be the line (v_t, v_1) . Let ϵ_i be $+1$ if the orientation of λ_i agrees with that of z ; -1 otherwise. Let r be the smallest integer such that for some $k \leq r-2$, λ_k and λ_r have a point in the same similarity class. Such an integer exists because λ_1 and λ_t have a point in the same similarity class. Then $T(\sum_{i=1}^r \epsilon_i \lambda_i)$ is a cycle y' of G' . By Lemma 1, there

² Without ambiguity, we shall also regard T as a mapping of chains.

exists a cycle y of G such that $T(y) = ky'$. The cycle y is *dep* B^* by definition of B^* , thus $\sum_{i=k}^{i=k'} \epsilon_i \lambda_i$ is *dep* B^* . Let $L_1 = z - \sum_{i=k}^{i=k'} \epsilon_i \lambda_i$. If L_1 is zero, then z is *dep* B^* . If L_1 is not zero, the process is continued as follows:

Let r' be the smallest integer such that for some k' satisfying $1 \leq k' < k$ or $r < k' \leq r' - 2$, $\lambda_{k'}$ and $\lambda_{r'}$ have a point in the same similarity class. Such an r' exists because λ_1 (or λ_{k+1} if L_1 does not contain λ_1) and λ_t (or λ_{r-1} if L_1 does not contain λ_t) have a point in the same similarity class. Let $L_2 = L_1 - \sum_{i=k'}^{i=r'} \epsilon_i \lambda_i$, where $\sum_{i=k'}^{i=r'} \epsilon_i \lambda_i$ is *dep* B^* as above. In the sequence of chains z, L_1, L_2, L_3, \dots , there must occur an L_j which is zero, showing that z is *dep* B^* .

We can now complete the proof of the theorem stated at the beginning of this section. Obviously no regular cycle of G can contain an exceptional line. Therefore the 1-1 correspondence between the cycles of G and those of G^0 preserves regularity. Thus for G^0 , we have $p^0 = p + k_e$, $k^0 = k$, $k_e^0 = 0$, $c^0 = c$, and $c_e^0 = c_e$.

Any dissimilarity cycle basis B is the union of its regular cycles B_{reg} and its exceptional cycles B_e . Using Lemma 2, one can show that B^* and B_e also constitute a dissimilarity cycle basis. Since c is an invariant of G , B^* and B_{reg} contain the same number of elements. Thus in G' , $p' = p + k_e$, $k' = k$, and $c' = c - c_e$. Now G' is connected, and in any connected graph, $\alpha_0 - \alpha_1 + \beta_1 = 1$ where $\alpha_0, \alpha_1, \beta_1$ are the numbers of points, lines, independent cycles respectively. Therefore $p' - k' + c' = 1$, and equation (2) follows immediately.

REFERENCES

1. F. Harary, *The number of graphs and directed graphs*, Bull. Amer. Math. Soc. Abstract 58-6-717.
2. F. Harary and R. Z. Norman, *The dissimilarity characteristic of Husimi trees*, Ann. of Math. vol. 58 (1953) pp. 134-141.
3. F. Harary and G. E. Uhlenbeck, *On the number of Husimi trees*, I, Proc. Nat. Acad. Sci. U.S.A. vol. 39 (1953) pp. 315-322.
4. D. König, *Theorie der endlichen und unendlichen Graphen*, Leipzig, 1936. Reprinted, New York, 1950.
5. R. Otter, *The number of trees*, Ann. of Math. vol. 49 (1948) pp. 583-599.
6. G. Pólya, *Kombinatorische Anzahlbestimmungen für Gruppen, Graphen und chemische Verbindungen*, Acta Math. vol. 68 (1937) pp. 145-254.
7. G. E. Uhlenbeck, *Some basic problems of statistical mechanics*, Gibbs Lecture, 1950.

UNIVERSITY OF MICHIGAN