Discrete Mathematics 40 (1982) 285–310
North-Holland Publishing Company

A CLASS OF SOLUTIONS TO THE GOSSIP PROBLEM, PART III

Douglas B. WEST

Department of Mathematics, Princeton University, Princeton, NJ 08544, USA

Received 13 August 1979
Revised 20 June 1981

We complete the study of NOHO-graphs, begun in Parts I and II of this paper. NOHO-graphs correspond to solutions to the gossip problem where *N*o *O*ne *H*ears his *O*wn information. These are graphs with a linear ordering on their edges such that an increasing path exists from each vertex to every other, but from no vertex to itself. We discard the two such graphs with no 2-valent vertices. In Part I, we translated these graphs into quadruples of integer sequences. In Part II, we characterized and enumerated the realizable quadruples and various subclasses of them. In Part III, we eliminate the overcounting of isomorphic graphs and obtain recurrence relations and generating functions to enumerate the non-isomorphic NOHO-graphs. If $u_m = (1, 1, 2, \ldots)$ satisfies $u_m = 3u_{m-1} - u_{m-3}$, then the number of non-isomorphic NOHO-graphs on $2m+2$ vertices is $\frac{1}{2}(u_m + u_{\lceil m/2\rceil+1} + u_{\lfloor m/2\rfloor+1} - u_{\lfloor m/2\rfloor})$ We also examine some related questions.

"Gossip is mischievous, light and easy
to raise, but grievous to bear and hard
to get rid of. No gossip ever dies
away entirely, if many people voice it;
it too is a kind of divinity."

Hesiod

1″. Introduction

In Part I of this paper [5], we described the history of the gossip problem. The original gossip problem asked for the minimum number of telephone calls between pairs of n gossips so that each will learn everyone else's information; for $n \geq 4$ the answer is $2n-4$. In Part I, we defined a new variation by adding the restriction that no gossip ever hears his own information. We gave a characterization of schemes that use the fewest calls among those schemes that transmit all the information without any gossip hearing his own. In Part II [6], we explored the properties of this characterization.

The purpose of this section is two-fold. First, we summarize the results of Parts I and II; second, we describe the results and methods used in the remainder of Part III. We will try to avoid unnecessary duplication; hence some of the results we cite may seem unmotivated. All the details, including definitions of standard terminology, can be found in the earlier parts. The numbers appended to various results in this summary are the labels they received previously; Sections 1–5 belong to Part I, 6–11 to Part II. If we need to refer to them subsequently, we will use the same labels.

Phrasing the gossip problem in graph terminology, we look for graphs whose edges are given a linear order such that there is an increasing path from each vertex x to every other vertex y. (We use $x \to y$ to mean "(an) increasing path from x to y".) Such linearly ordered edge sets are called *pooling schemes*. If there is no $x \to x$ for any vertex x, the graph (with edge ordering) *satisfies NOHO*. Pooling schemes on n vertices can satisfy NOHO precisely when n is even, and the smallest such graphs have $2n-4$ edges (Lemma 2.4). Our purpose in Part III is to enumerate the non-isomorphic 'NOHO-graphs' on n vertices. *NOHO-graphs* on $n \geq 4$ vertices are those having a 2-valent vertex and having $2n-4$ edges that can be ordered to obtain a pooling scheme satisfying NOHO. The 2-valent vertex condition eliminates two aberrant graphs.

NOHO prohibits triangles (Remark 2.2). Given an ordering of the edges, we use $f(x)$ to denote the first neighbor of vertex x and $l(x)$ to denote its last neighbor. $F(G)$ is the collection of first-edges $(x, f(x))$; similarly $L(G)$ denotes the collection of last-edges. $F(G)$ and $L(G)$ each forms a complete matching in a NOHO-graph, and they are disjoint (Lemma 2.3). The remaining $n-4$ edges, called $M(G)$, produce two isolated vertices and two caterpillars on $\frac{1}{2}n-1$ vertices each (Lemma 4.2). (A caterpillar is a tree with a path that contains or neighbors every vertex.) Henceforth, we will set $m=\frac{1}{2}n-1$, since this is a natural parameter. In drawing NOHO-graphs, we adopt the convention of drawing edges of $F(G)$ as dotted lines, edges of $L(G)$ as dashed lines, and edges of $M(G)$ as solid lines.

The main result of Part I is a description of NOHO-graphs by quadruples of integer sequences. P and Q are permutations which describe the placement of first and last edges, respectively. S and T are binary sequences which describe the placement of edges within the caterpillars remaining. The sequences refer to particular vertex labels.

Given an ordering of the edges in a NOHO-graph, define the *canonical numbering* of its vertices to be a labelling in which the vertices receive the labels $\{x_j^i\colon i=1,2;\ j=0,1,\ldots,m\}$, distributed as follows. Assign x_0^i to the 2-valent vertices. Let C^i be the caterpillars of $M(G)$. Each is a tree of increasing paths out of one vertex and into another. This gives us a well-defined labelling of the vertices of C^i as $\{x_j^i\}$ so that x_j^i is the jth to receive information from x_0^i. We also refer to x_0^i as x_{m+1}^{3-i}, and we use $\bar{C}^i$ to denote the 'extended' caterpillar that consists of x_0^i, C^i, and x_{m+1}^i. The increasing path from x_0^i to x_{m+1}^i in $\bar{C}^i$ is called its *distinguished path*.

The canonical numbering summarizes which increasing paths and adjacencies occur in the caterpillars. These properties (Remark 4.3) are listed below. (In addition to the notation $x \to y$, we use $x \sim y$ to mean "x is adjacent to y" and use $x \not\sim y$ for non-adjacency.)

(a) $\bar{C}^i$ contains $x_j^i \to x_k^i$ if and only if $j<k$ or $x_j^i \sim x_k^i$.

(b) x_k^i neighbors exactly one x_j^i such that $j<k$.

(c) If x^i_k neighbors any x^i_r with $r>k$, then it neighbors every x^i_j with $k<j<r$.

Given any caterpillar that is a tree of increasing paths out of one vertex and into another, there is an associated binary sequence that determines the arrangement of the edges (Remark 4.4). The canonical numbering conforms to this in such a way that in the resulting sequences S and T describing the two caterpillars, $S_j=1$ if and only if x^1_j lies on the distinguished path in $\bar{C}^1$, and $T_j=1$ if and only if x^2_{m+2-j} lies on the distinguished path in $\bar{C}^2$ (Remark 4.5). These sequences uniquely determine the labelled caterpillars. The first and $m+1$st elements of S and T are always 1 and are often dropped.

Due to NOHO and Remark 4.3, the first and last neighbors of a vertex in C^i must belong to $\bar{C}^j$. Hence we define $P_r=s$ if $f(x^1_r)=x^2_s$ and $Q_r=s$ if $l(x^1_r)=x^2_s$. P and Q are permutations; P on $\{1,\ldots,m+1\}$, Q on $\{0,\ldots,m\}$. These determine $F(G)$ and $L(G)$.

Hence, any NOHO-graph G is uniquely determined by $(P(G); Q(G); S(G); T(G))$ (Theorem 4.6). A quadruple that can arise as the defining quadruple of a NOHO-graph with respect to some ordering is called a *realizable quadruple*. The remainder of Part I was devoted to showing constructively that NOHO-graphs are Hamiltonian (Theorem 5.2), bipartite (Lemma 5.3), and planar (Theorem 5.4). In fact, we will see that it has a planar representation in which all the faces are 4-cycles. These results use the fact (Lemma 5.1) that, if a path alternating between first-edges and last-edges is grown from x^1_0, then the indices in C^1 increase and those in C^2 decrease, so in a sense the edges do not 'cross'.

In Part II, we characterized and enumerated realizable quadruples. The major step here was that any pair of sequences in a realizable quadruple uniquely determines the remaining pair (Theorem 7.3). For this, it is necessary to determine the form of a realizable P and relations between $P(G)$, $Q(G)$, and $S(G)$. In particular, P is composed of substrings that have the form $(r, r+s, r+s-1,\ldots,r+1)$ or consist of a single element that exceeds all subsequent elements (Lemma 6.5). This result uses Remark 4.3 and the edges implied by increasing pairs in P; if $P_i<P_j$ with $i<j$, then $x^1_i\sim x^1_j$ and $x^2_{P_i}\sim x^2_{P_j}$ (Lemma 6.3). These substrings are the 'reversions' of P, where the *reversions* of a permutation are the maximal substrings in which the first element is the least.

For relating $P(G)$, $Q(G)$, and $S(G)$ and proving most of the structural results about NOHO-graphs, the basic result is Lemma 7.1. This says that $S_i=0$ if and only if $P_{i+1}=Q_i$ and $S_i=1$ if and only if $P_{i+1}=Q_{b(i)}$ where $b(i)$ is the index of the previous 1 in S. There is a corresponding result for P, Q, T. Using slightly different notation, Lemma 7.1* says $T_j=0$ if and only if $f(x^2_{m+2-j+1})=l(x^2_{m+2-j})$ and $T_j=1$ if and only if $f(x^2_{m+2-j+1})=l(x^2_{m+2-c(j)})$ where $c(j)$ is the index of the next 1 in S.

These can be interpreted graphically. Remarks 4.3 and 4.5 imply $x^1_i\sim x^1_{b(i)}$ and $x^2_{m+2-j}\sim x^2_{m+2-c(j)}$. Therefore, Lemmas 7.1 and 7.1* imply that the first and last neighbors of any point in C^i are joined by a path of two edges in the opposite (extended) caterpillar (Corollary 7.2).

A set of four necessary conditions developed in Section 6 turns out to be sufficient for a quadruple to be realizable (Theorem 7.4). We will not need those conditions. In Part III we will concentrate on the pair $(S(G); T(G))$. Since any pair determines the remaining pair, we will need only the necessary and sufficient conditions for the realizability of $(S; T)$.

To develop those conditions we defined concatenation of NOHO-graphs. Given two NOHO-graphs G_1 and G_2 with canonical numberings $\{x_j^i\}$ and $\{y_j^i\}$, the *concatenation* $G_1 + G_2$ is a graph obtained by deleting $\{x_0^2, y_0^1\}$ and their incident edges and identifying y_1^1 with $l(x_0^2)$ and x_1^2 with $l(y_0^1)$. This graph has $n_3 = n_1 + n_2 - 4$ vertices and $2n_3 - 4$ edges. To obtain the appropriate edge-ordering on $G_1 + G_2$, the edge-orderings on G_1 and G_2 are merged naturally. The first or last edges are those that were first or last edges originally. The edges that were in $C^1(G_1)$ appear before those that were in $C^1(G_2)$, and those of $C^2(G_2)$ appear before those of $C^2(G_1)$. With this ordering, $G_1 + G_2$ is a NOHO-graph (Lemma 9.1).

Writing the NOHO-graph determined by $(S; T)$ as $G(S; T)$, the effect of concatenation is easily expressed (Remark 9.2). If $G_1 = G(S_2, \ldots, S_j; T_2, \ldots, T_j)$ and $G_2 = G(S_2', \ldots, S_k'; T_2', \ldots, T_k')$, then

$$G_1 + G_2 = G(S_2, \ldots, S_j, S_2', \ldots, S_k'; T_2, \ldots, T_j, T_2', \ldots, T_k').$$

Finally, deconcatenation also makes sense. If $(S_2, \ldots, S_m; T_2, \ldots, T_m)$ is realizable, then an initial segment $(S_2, \ldots, S_k; T_2, \ldots, T_k)$ is realizable if and only if the final segment $(S_{k+1}, \ldots, S_m; T_{k+1}, \ldots, T_m)$ is also realizable (Lemma 9.3).

Irreducible quadruples are those that cannot be obtained by concatenating smaller quadruples. Every realizable quadruple has a unique decomposition as a concatenation of irreducible quadruples (Lemma 10.1). These are called its irreducible components. Thus it suffices to determine the pairs $(S; T)$ which determine irreducible quadruples. The necessary and sufficient conditions (Lemma 11.5), which we use habitually in Part III, are as follows (the *weight* of a binary sequence is the number of 1's it contains):

(a) $\{S_2, T_2\} = \{S_m, T_m\} = \{0, 1\}$,

(b) $S_k = T_k$ for $2 < k < m$,

(c) $(S_2, \ldots, S_m)$ and $(T_2, \ldots, T_m)$ have odd weight.

The major structural result leading to this characterization will be useful also in Part III. This characterization, Lemma 11.1, describes the placement of edges in an irreducible $G(S; T)$. If $\{S_2, T_2\} = \{0, 1\}$ and $S_k = T_k$ for $2 < k \leq r$, then $S_2 = 0$ implies $x_r^1 \sim x_{m+2-r}^2$, while $S_2 = 1$ implies $x_r^1 \sim x_{m+2-b(r)}^2$. If $(S_2, \ldots, S_{r-1})$ has even weight, the indicated edge belongs to $F(G)$; otherwise it belongs to $L(G)$.

Earlier in Part II, we enumerated symmetric and reversible quadruples. The operation of *reflection* (through the 'center') is a vertex permutation that interchanges x_j^1 and x_j^2. If the new labelling gives rise to the same quadruple, then the graph and quadruple are *symmetric*. A realizable $(S; T)$ generates a symmetric quadruple if and only if $S_k = T_{m+2-k}$ for all k.

Reversing is the operation of reversing the edge-ordering on a NOHO-graph. If the quadruple that arises from the canonical numberings associated with the orderings are the same, then the graph and quadruple are *reversible*. The graph generated by a realizable $(S;T)$ is reversible if and only if $(S_k;T_k) = (S_{m+2-k};T_{m+2-k})$ for all k.

As mentioned earlier, our task here is to enumerate the non-isomorphic NOHO-graphs. This means we must determine the ways in which different quadruples can arise from a single NOHO-graph, depending on what legal edge-ordering is associating with it. More specifically, which pairs $(S;T)$ will generate the same graph, besides the pairs obtained by reflecting and reversing?

In Section 12, we study the question of twisting a pair $(S;T)$. Applying a *twist* between k and $k+1$ means interchanging S_j and T_j for $j \leq k$ or for $j > k$. The main result here is that a twist produces an isomorphic NOHO-graph if and only if the twist occurs next to an irreducible component in $(S;T)$ of length 1.

In Sections 13 and 14 we show that these are the *only* ways of obtaining isomorphic NOHO-graphs. Section 13 contains a detailed study of the adjacencies of vertices along the distinguished paths in a NOHO-graph according to their distance from x_0^1. These characteristics are invariant under isomorphism, except for the possibility of mapping x_0^1 to x_0^2. Section 14 contains the main theorem: the set of distances and degrees of these vertices determines $(S;T)$ up to twisting and reversing.

In Section 15, we apply this to enumerate the non-isomorphic NOHO-graphs. We obtain a recurrence relation for the number of inequivalent pairs $(S;T)$ under twisting. On $2m+2$ vertices, this is $u_m = 3u_{m-1} - u_{m-3}$. This produces somewhat less than the 3^{m-2} realizable quadruples obtained in Part II. To eliminate the remaining overcounting due to reversing, it is necessary to add the correctly counted symmetric or irreducible NOHO-graphs and divide by 2. Thus the number of non-isomorphic NOHO-graphs on $2m+2$ vertices is $\frac{1}{2}(u_m + u_{\lceil m/2 \rceil + 1} + u_{\lfloor m/2 \rfloor + 1} - u_{\lfloor m/2 \rfloor})$.

Finally, in Section 16 we discuss some related gossip questions.

12. Twisted NOHO-graphs

We say two realizable quadruples are *equivalent* if they are both realized by the same graph. The operations of reflecting and reversing generate a four element group that acts on realizable quadruples to produce equivalence classes. If these were the only ways to obtain more than one quadruple from a graph, we could now count the number of non-isomorphic NOHO-graphs by Polya's theorem. However, there is yet another wrinkle in this problem; sometimes we can 'twist' $(S;T)$ without changing the graph. For example, the underlying graphs generated by (01110; 10101) and (01101; 10110) are isomorphic, as shown by the labellings in Fig. 12.1.

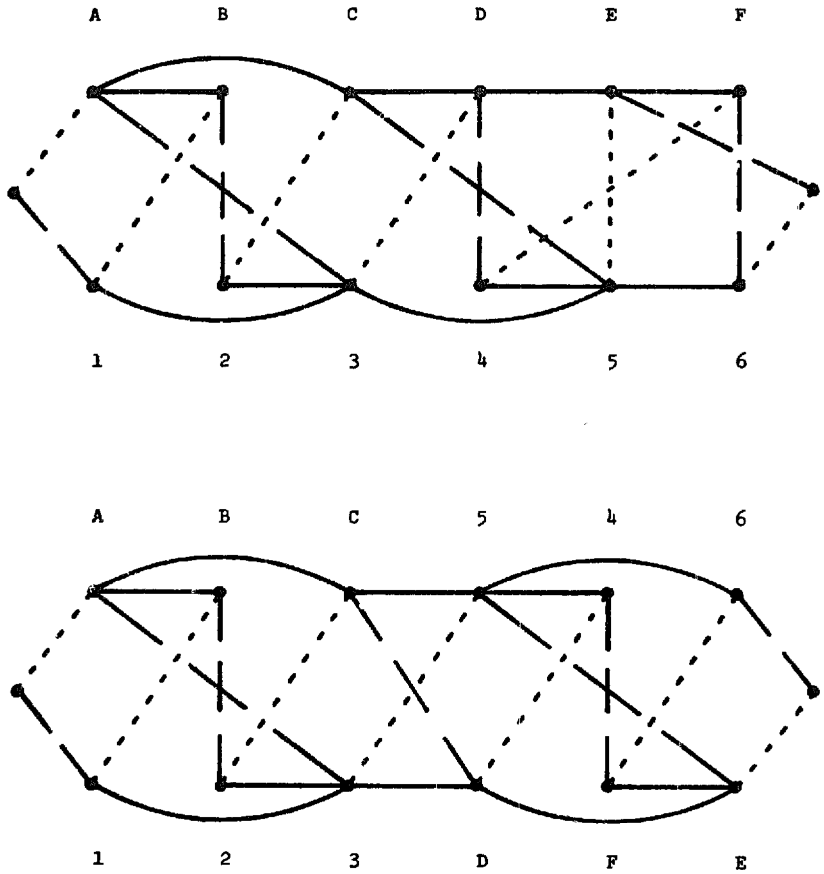

Fig. 12.1. Realizing (01110; 10101) and (01101; 10110).

We define *twisting between k and* $k+1$ for a pair of sequences $(S; T)$ to mean interchanging S_i with T_i for $i > k$, or interchanging S_i with T_i for $i \leq k$. To distinguish these, we say the former is a twist *after* k, and the latter is a twist *before* $k+1$. We want to know when twists produce realizable quadruples and, more specifically, when they produce new quadruples realizable by the same graph. Since reversing and reflecting are always available, whether the quadruples are equivalent will not depend on whether the twist occurs before $k+1$ or after k. Note that twisting after 1 or before $m+1$ corresponds to performing both reflecting and reversing. In Sections 13–14 we will show that twisting is the only other operation we need to get all the equivalence relations between quadruples.

In testing isomorphism, information about the degrees of vertices and their distances from x_0^1 will be helpful. The former is all we need to analyze twisting. For the purposes of the next remark, and in general throughout this section, S and T are treated as having $m+1$ elements by including $(S_1; T_1) = (S_{m+1}; T_{m+1}) = (1; 1)$. The next result follows directly from the canonical numbering (Remark 4.3).

Remark 12.1. *For a NOHO-graph, the degrees of vertices in the caterpillars are*

$$x_i^1:\begin{cases}3, & \text{if } S_i=0,\\ 3+j-i, & \text{if } S_i=1 \text{ and } S_j \text{ is the next } 1 \text{ in } S.\end{cases}$$

$$x_i^2:\begin{cases}3, & \text{if } T_{m+2-i}=0,\\ 3+j-i, & \text{if } T_{m+2-i}=1 \text{ and } T_{m+2-j} \text{ is the preceding } 1 \text{ in } T.\end{cases}$$

Lemma 12.2. *A twist within an irreducible component of a realizable $(S;T)$ produces an unrealizable $(S';T')$.*

Proof. The initial irreducible components are unchanged. The later irreducible components have been reflected and reversed, hence are still realizable. By Lemma 9.3, realizability requires realizability of the new component where the twist occurred. Since $(S;T)$ for an irreducible quadruple differ only in positions 2 and m, the only change here is the interchange of 0 and 1 in one of those positions. This turns the component into the even-weight type forbidden by Lemma 11.5. □

Lemma 12.3. *A twist between two irreducible components of length at least two in a realizable $(S;T)$ produces a realizable but inequivalent $(S';T')$.*

Proof. $(S';T')$ is clearly realizable, since it is a concatenation of realizable segments. However, examining the degrees of the vertices will expose a difference.

Without loss of generality, we may assume that the twist occurs after k and that $(S_k;T_k)=(S_{k+1};T_{k+1})=(1;0)$. Let $d(x)$ and $d'(x)$ be the degree of x in the graphs generated by $(S;T)$ and $(S';T')$, and consider the computation of degrees according to Remark 12.1. For $j\leqslant k$, $d(x_j^1)=d'(x_j^1)$ and $d(x_{m+2-j}^2)=d'(x_{m+2-j}^2)$, except that $d(x_k^1)\neq d'(x_k^1)$. For $j>k$, $d(x_j^1)=d'(x_{m+2-j}^2)$ and $d(x_{m+2-j}^2)=d'(x_j^1)$, except that $d(x_{m+2-c(k)}^2)\neq d'(x_{c(k)}^1)$. So, we need only show $\{d(x_k^1), d(x_{m+2-c(k)}^2)\}\neq\{d'(x_k^1), d'(x_{c(k)}^1)\}$. Since the next 1 in S follows S_k immediately, $d(x_k^1)=4$. Both of $d'(x_k^1)$ and $d'(x_{c(k)}^1)$ exceed 4, since $S'_{k+1}=0$ and $T'_k=0$ serve as intervening zeros. □

Theorem 12.4. *An equivalent quadruple can be obtained by applying a twist to a realizable $(S;T)$ if and only if the twist is applied before or after an irreducible component $(S;T)=(1;1)$ of length 1.*

Proof. Given the preceding results, we need only exhibit a vertex mapping π to verify the isomorphism in the case described. Let $(S';T')$ be the new pair of sequences; we map the new graph into the old. Without loss of generality, we may assume that the twist occurs after k, that $(S_k;T_k)=(1;0)$, and that $S'_{k+1}=T'_{k+1}=T'_{k+2}=1$. In Fig. 12.1, such a twist has occurred with $k=3$.

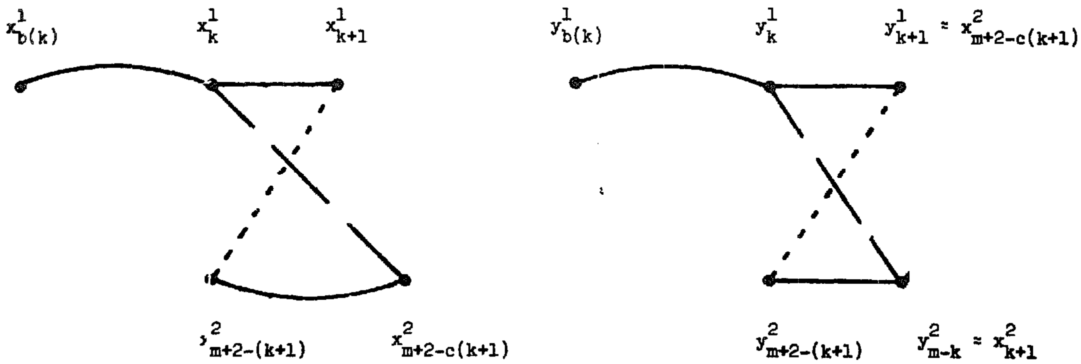

Fig. 12.2. Relabelling after a legal twist.

Let $\{y^i_j\}$ and $\{x^i_j\}$ be the vertices of $G(S'; T')$ and $G(S; T)$ respectively, with the canonical numbering. For $j \leq k$, let $\pi(y^1_j) = x^1_j$ and $\pi(y^2_{m+2-j}) = x^2_{m+2-j}$; also put $\pi(y^2_{m+2-(k+1)}) = x^2_{m+2-(k+1)}$. The initial segment is the same in both pairs and realizable by adding a vertex that neighbors y^1_k and $y^2_{m+2-(k+1)}$, so the vertices mapped above are adjacent if and only if their images are adjacent.

In both $(S; T)$ and $(S'; T')$, consider what happens when we concatenate the initial segment in positions 2 through k with the final segment in positions $k+1$ through m. The edges guaranteed by the definition of concatenation and the presence of the 1's assumed in $(S; T)$ and $(S'; T')$ tell us that both $(x^1_k, x^2_{m+2-c(k+1)}, x^2_{m+2-(k+1)}, x^1_{k+1})$ and $(y^1_k, y^2_{m-k}, y^2_{m+1-k}, y^1_{k+1})$ are four-cycles. (See Fig. 12.2) We can let $\pi(y^1_{k+1}) = x^2_{m+2-c(k+1)}$ and $\pi(y^2_{m-k} = x^2_{k+1})$ to preserve image adjacencies for y^1_{k+1} and y^2_{m-k} among the vertices mapped thus far.

Now shift the bits $(1; 1)$ in position $k+1$ from the final segment to the initial segment. Still both final segments are realizable. The twist has the effect of reversing the edge ordering and interchanging the caterpillars in the graphs H and H' realizing the final segments. H and H' are isomorphic. Their 2-valent vertices to be eliminated for concatenation correspond to x^1_k (or x^2_{m+1-k}) and y^1_k (or y^2_{m+1-k}). The vertices adjacent to them are $\{x^1_{k+1}, x^2_{m+2-c(k+1)}\}$ and $\{y^1_{k+1}, y^2_{m-k}\}$. The switch made by π is the same as the switch made by the isomorphism between H' and H obtained by reflecting and reversing. So, π can be completed as desired by using that isomorphism. □

13. Distances from x^1_0

To complete our study of isomorphism, we want to show that the only ways to obtain equivalent quadruples are those we have described. Reversing is always available, so we can assume that a candidate isomorphism between $G(S'; T')$ and $G(S; T)$ takes y^1_0 to x^1_0. Therefore, we examine the relationship between a graph and the quadruples it realizes in terms of the distance of vertices from x^1_0.

Henceforth, let $\delta(x_j^i)$ be the number of edges in the shortest path (not necessarily an increasing path) between x_j^i and x_0^1. We call this the *x-distance* of x_j^i. We will show that the x-distances uniquely determine the irreducible NOHO-graphs. We will need detailed knowledge about the placement of certain edges.

Lemma 13.1. *Suppose $G(S; T)$ is an irreducible NOHO-graph and $S_k = 1$ for some k with $2 < k < m$. Let vertices x_j^1 and x_{m+2-j}^2 with $j > k$ be "right-vertices", and let those with $j < k$ be "left-vertices". Then there is only one edge joining a right-vertex and a left-vertex. This edge is a last-edge. If $(S_2, \ldots, S_{k-1})$ has even weight, it is the last call of $x_{m+2-c(k)}^2$, otherwise it is the last call of $x_{b(k)}^1$. When $S_m = 1$, the same applies, except that caterpillar edges in C^2 may join x_1^2 to left-vertices.*

Proof. First note that no caterpillar edges can cross from left to right, which is immediate from $S_k = T_k = 1$ and the canonical numbering. For first and last edges, the two cases are illustrated in Fig. 13.1; we consider them in parallel. By Lemma 11.1 we know $x_{b(k)}^1 \sim x_{m+2-k}^2$ and $x_k^1 \sim x_{m+2-c(k)}^2$. If $(S_2, \ldots, S_{k-1})$ has even weight, these edges lie in $L(G)$ and $F(G)$, respectively, else vice versa. The argument below includes the two subcases $b(k) = k-1$ and $c(k) = k+1$, which are not drawn in Fig. 13.1. Those cases do appear graphically in Figs. 13.2 and 13.3.

In the even case, Lemma 7.1 gives $f(x_{m+2-k}^2) = x_{k+1}^1$, while in the odd case its reflection (Lemma 7.1*) gives $f(x_k^1) = x_{m+3-k}^2$. In either case, a first-edge from left to right would cross this and generate an increasing pair in P. By Lemma 6.3(a), an increasing pair in P forces the presence of edges in the caterpillars joining the endpoints of the two first-edges that cross. However, this would give x_{k+1}^1 or x_{m+3-k}^2 two preceding neighbors in its caterpillar, which the canonical numbering (Lemma 4.3(b)) prohibits.

So, only last-edges can join left and right vertices. From the path-growing by Lemma 7.1 or 7.1*, it follows that in the even or odd case $(x_{m+2-c(k)}^2, l(x_{m+2-c(k)}^2))$ or $(x_{b(k)}^1, l(x_{b(k)}^1))$ is a last-edge from left to right. No other last-edge can cross this one, since in all cases Lemma 6.3 would demand an edge which cannot be added

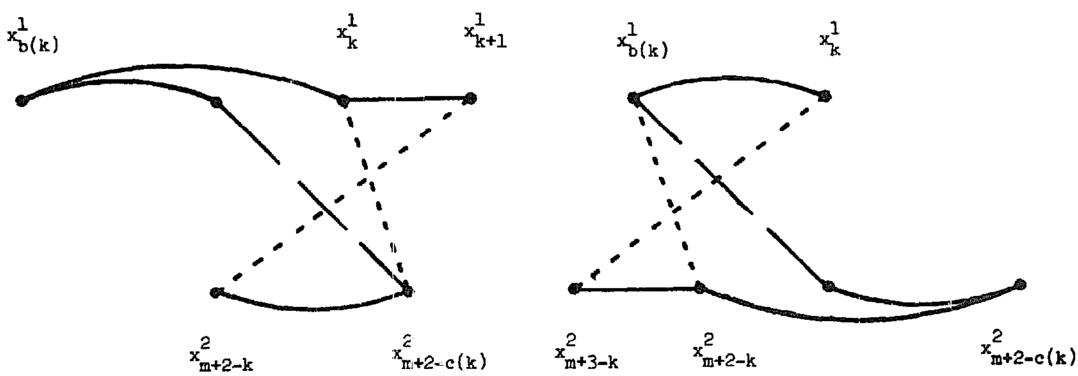

Fig. 13.1. Placement of edges in irreducible NOHO-graphs; $(S_2, \ldots, S_{k-1})$ has even or odd weight.

to these caterpillars without violating the caterpillar numbering. On the other hand, no last-edge can be fit in without crossing this one. Adding the first-edge incident to $l(x^2_{m+2-c(k)})$ or $l(x^1_{b(k)})$ in Fig. 13.1, it is easy to see that such an edge would cross a first-edge from both alternating paths, and hence could not belong to either, by Lemma 5.1. □

The next remark, which follows from the fact that NOHO-graphs are bipartite (Lemma 5.3), will be used implicitly in the discussion of x-distance.

Remark 13.2. *The x-distances of adjacent vertices in a NOHO-graph differ by* 1.

The x-distance usually increases by 1 with each step along either caterpillar path from x^1_0 towards x^2_0. However, it is sometimes possible to short-circuit four steps in a caterpillar path by crossing to the other caterpillar, taking one step there, and crossing back. If $S_k = 1$ or $T_k = 1$ and $\delta(x^1_k) = \delta(x^1_{b(k)}) - 1$ or $\delta(x^2_{m+2-k}) = \delta(x^1_{m+2-b(k)}) - 1$, we say a *descent* occurs *at* k. If the smaller x-distance is r when a descent occurs at k, we call it a descent *to* r. If the vertex mentioned in the previous sentence is in C^i, we say the descent is *in* C^i.

Before embarking on the next two lemmas, which are rather technical, it is appropriate to describe what will be achieved by examining x-distances and descents. Lemma 13.3 establishes necessary conditions for descents. More importantly, it points out that not many vertices can have the same x-distance. In particular, when we look at this in more detail in Corollary 13.5, we will see that there can be at most one descent to r, and at most three r-vertices with degree more than three.

The arrangement of these 'high-valent' r-vertices is intimately related to the structure of irreducible NOHO-graphs. In fact, they will enable us to reconstruct such a graph given only the degrees and x-distances of the vertices. Lemma 13.4 takes the first step in this direction, expanding Lemma 13.3 into necessary and sufficient conditions for the occurrence of descents. Theorem 14.1 completes this process by determining when irreducible quadruples can give rise to the same sets of (degree, x-distance) pairs for the vertices.

With this program in mind, we proceed to the details. The reader may find the statement and proof of Lemma 13.3 easier to follow by referring to Figs. 13.2 and 13.3. Looking at these figures, it is interesting to note that this is our first lemma that deals simultaneously with a cross-section of both alternating paths. The basic tools in the proof are the structural Lemma 13.1 and the fact that a vertex at distance r from x^1_0 must neighbor a vertex with x-distance $r-1$. Henceforth, for compactness, we call a vertex with x-distance r an *r-vertex*. Throughout the remainder of this section, we will also use implicitly the $(S; T)$ characterization of NOHO-graphs (Lemma 11.5). In particular, $S_k = T_k$ for $2 < k < m$, so $c(b(k)) = k$, etc,, as needed.

Lemma 13.3. *If a descent to r occurs at k in an irreducible NOHO-graph, then $(S_2, \ldots, S_{k-1})$ has odd weight. If the descent occurs in C^1, then*

(a) $S_{k-1} = 1$,

(b) $\delta(x^2_{m+2-k}) = r$, *and*

(c) *all right-vertices v have $\delta(v) \geq r+2$ except for neighbors of $\{x^1_k, x^2_{m+2-k}\}$.*

If the descent occurs in C^2, then

(a) $S_{b(k)-1} = 1$,

(b) $\delta(x^1_k) = r+2$, *and*

(c) *all right-vertices have $\delta(v) \geq r+2$ except $x^2_{m+2-c(k)}$.*

Proof. Consider a path from a right-vertex to x^1_0. It must cross from right-vertices to left-vertices at some point, so by Lemma 13.1 it must pass through one of three vertices. These are $\{x^1_k, x^2_{m+2-k}, x^1_{b(k)}\}$ if $(S_2, \ldots, S_{k-1})$ has odd weight, or $\{x^1_k, x^2_{m+2-k}, x^2_{m+2-c(k)}\}$ if $(S_2, \ldots, S_{k-1})$ has even weight. Consider the x-distances of these vertices.

First suppose the descent is in C^1, so $\delta(x^1_k) = r$ and $\delta(x^1_{b(k)}) = r+1$. By Lemma 11.1, $x^1_{b(k)} \sim x^2_{m+2-k}$, so $\delta(x^2_{m+2-k})$ is r or $r+2$. Thus, all the right-vertices have x-distance as large as claimed. Fig. 13.2 illustrates the cases of odd or even weight in $(S_2, \ldots, S_{k-1})$; consider the odd case first. We must find an $(r-1)$-vertex to neighbor x^1_k; it must be a left vertex. $x^1_{b(k)}$ doesn't help, but there is another vertex available. Since $(S_2, \ldots, S_{k-1})$ has odd weight, $(x^1_k, x^2_{m+2-c(k)}) \in L(G)$. Applying Lemma 7.1* to x^2_{m+2-k}, we have $(x^1_k, x^2_{m+3-k}) \in F(G)$. No other left vertex neighbors x^1_k, so $\delta(x^2_{m+3-k}) = r-1$. Now $\delta(x^2_{m+2-k}) = r$. Finally, to prove (a), suppose $S_{k-1} = 0$. Then x^2_{m+3-k} neighbors only three vertices, by Remark 12.1. These are $\{x^2_{m+2-k}, x^1_k, x^1_{k-1}\}$. By applying Lemma 7.1 to $S_{k-1} = 0$, we have $x^1_{k-1} = l(x^2_{m+3-k})$. Since $\delta(x^2_{m+3-k}) = r-1$ and $\delta(x^1_{b(k)}) = r+1$ for neighbors of x^1_{k-1}, we have $\delta(x^1_{k-1}) = r$. But now x^2_{m+3-k} neighbors only r-vertices and can't have x-distance $r-1$. Therefore, $S_{k-1} = 1$ and $b(k) = k-1$.

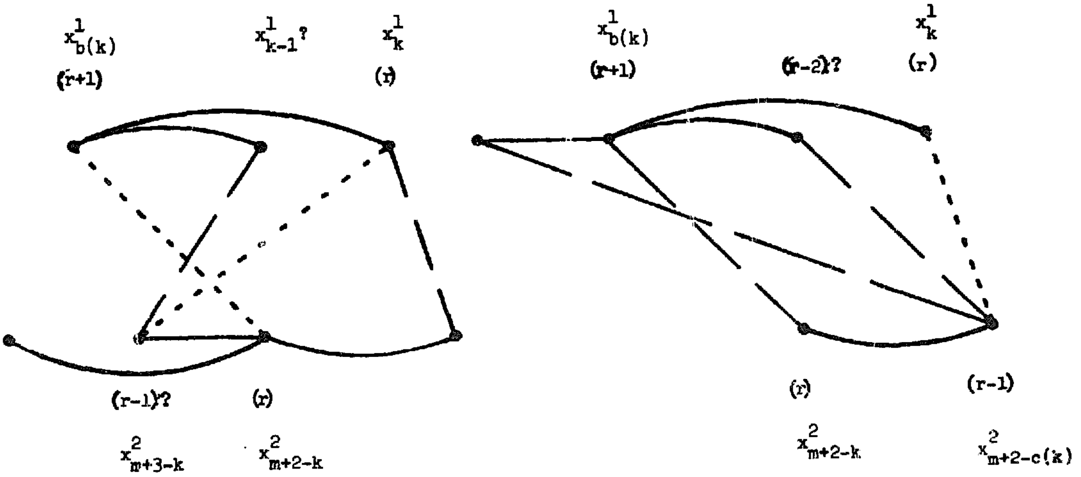

Fig. 13.2. Descents in C^1; $(S_2, \ldots, S_{k-1})$ has odd or even weight.

In the even case, we cannot find an $(r-1)$-vertex to neighbor x_k^1. This time x_k^1 neighbors no left-vertex except $x_{b(k)}^1$, but $f(x_k^1)=x_{m+2-c(k)}^2$. Again, $f(x_k^1)$ is the only possible neighboring $(r-1)$-vertex. Recall that $(x_{m+2-c(k)}^2, l(x_{m+2-c(k)}^2))$ is the only left-right edge, by Lemma 13.1. By Corollary 7.2 the first and last neighbors of $x_{m+2-c(k)}^2$ must be joined by a path of two edges in $\bar{C}^1$. The only possible intermediate vertex is $x_{b(k)}^1$. But now $\delta(l(x_{m+2-c(k)}^2))=\delta(x_{m+2-k}^2)=r$, and $x_{m+2-c(k)}^2$ neighbors no $(r-2)$-vertex. Now consider a descent at k in C^2. First note that $b(k)\geq 2$. If $(S_2;T_2)=(0;1)$, then $x_m^2 \sim x_0^1$, but a descent in C^2 at $c(2)$ would require $\delta(x_m^2)>1$. Hence $b(k)=c^{-1}(k)$ and it makes sense to say that the $(r+1)$-vertex in C^2 for this descent is $x_{m+2-b(k)}^2$. We can proceed much as before once we show that $\delta(x_{b(k)}^1)=r+1$. It must be $r-1$ or $r+1$, since, by Lemma 11.1, $x_{b(k)}^1 \sim x_{m+2-k}^2$. Similarly, $x_{b^2(k)}^1 \sim x_{m+2-b(k)}^2$, so $\delta(x_{b^2(k)}^1)\geq r$. If $\delta(x_{b(k)}^1)=r-1$, then there is a descent to $r-1$ at $b(k)$ in C^1. By fact (b) about such descents, this would require $\delta(x_{m+2-b(k)}^2)=r-1$, which is false, so $\delta(x_{b(k)}^1)=r+1$.

Suppose $\delta(x_k^1)=r$; then there is a descent at k in C^1. By the preceding analysis of such descents, $b(k)=k-1$. Also, the only neighbor of x_k^1 which can be an $(r-1)$-vertex is $x_{m+3-k}^2=f(x_k^1)$. But if $b(k)=k-1$, then this means $\delta(x_{m+2-b(k)}^2)=r-1$, eliminating the descent in C^2.

Therefore, we have $\delta(x_{b(k)}^1)=r+1$ and $\delta(x_k^1)=r+2$. Furthermore, $\delta(x_{m+2-c(k)}^2)=r+1$, since that vertex neighbors both x_k^1 and x_{m+2-k}^2. In the odd or even case, the x-distances of the three crucial vertices have now been determined; all vertices on the right have x-distance at least $r+2$, except neighbors of x_{m+2-k}^2. Now, how can x_{m+2-k}^2 neighbor an $(r-1)$-vertex? The two cases are illustrated in Fig. 13.3. In the even case, $l(x_{m+2-k}^2)=x_{b(k)}^1$ and $f(x_{m+2-k}^2)=x_{k+1}^1$. Therefore, an $(r-1)$-vertex neighboring x_{m+2-k}^2 can only be x_{m+2-j}^2, for some $b(k)<j<k$. However, these vertices neighbor only vertices which neighbor $x_{b(k)}^1$, so they can't neighbor any $(r-2)$-vertices.

In the odd case, the same argument holds against x_{m+2-j}^2. However, now $x_{b(k)}^1=f(x_{m+2-k}^2)$, and $l(x_{m+2-k}^2)$ can be the desired $(r-1)$-vertex neighbor of x_{m+2-k}^2. Here the argument parallels that for descents in C^1. If $S_{b(k)-1}=0$, then $l(x_{m+2-k}^2)=x_{b(k)-1}^1$, by Lemma 7.1. This time the three neighbors of this vertex—$\{x_{b^2(k)}^1, x_{m+3-b(k)}^2, x_{m+2-k}^2\}$—all neighbor $x_{m+2-b(k)}^2$, so again none of them are $(r-2)$-vertices. So, if there is a descent in C^2, it must be in the odd case and have $S_{b(k)-1}=1$. □

Let $\beta(k)$ be the index of the last descent preceding k. Set $\beta(k)=1$ if there is none. The next lemma enables us to read off x-distances from $(S;T)$.

Lemma 13.4. *Suppose $G(S;T)$ is an irreducible NOHO-graph. The first descent in G occurs in C^1 if $S_2=1$ and in C^2 if $T_2=1$; thereafter they alternate between the caterpillars. For a descent in C^1 set $q=k$ and $t=2$; for a descent in C^2 set $q=b(k)$ and $t=4$. Then a descent occurs at k if and only if $b(q)=q-1$ and $(S_{\beta(k)},\ldots,S_{k-1})$ has even weight at least t. However, $t\geq 4$ is required for the first descent even when it is in C^1.*

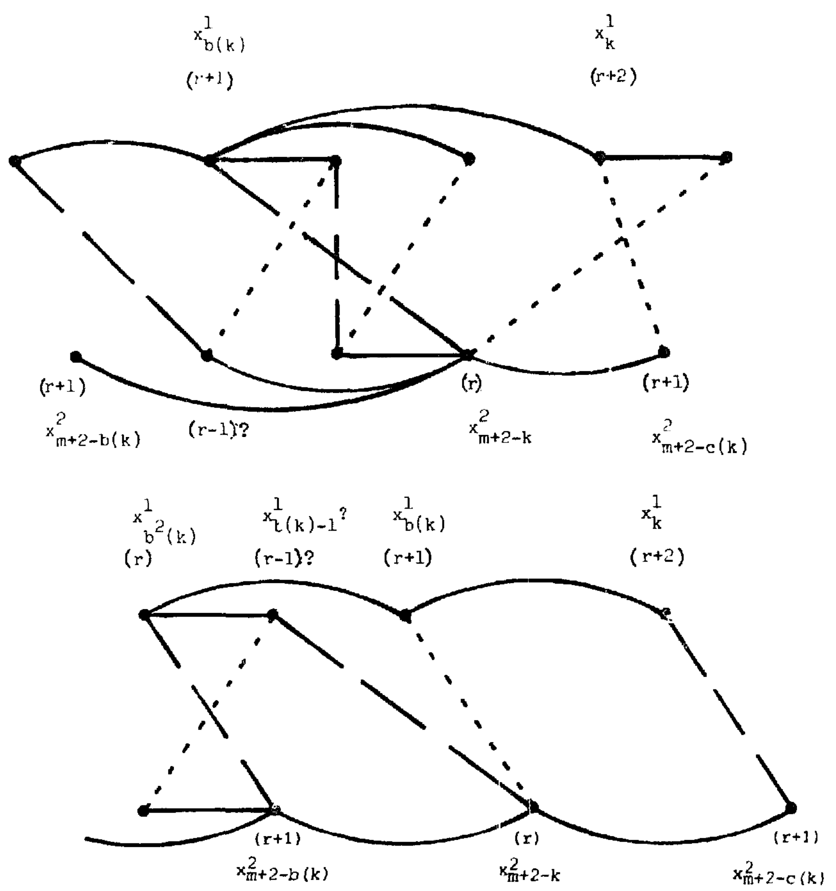

Fig. 13.3. The even and odd cases for a descent in C^2.

Proof. Lemma 13.3(b) gave the relationships between x-distances of vertices when descents occur. In particular, $\delta(x^1_k)=\delta(x^2_{m+2-k})$ for a descent at k in C^1, and $\delta(x^1_k)=\delta(x^2_{m+2-k})+2$ for a descent at k in C^2. The vertices of the distinguished paths in the caterpillars correspond to $S_j=T_j=1$. Since they do form paths, by Remark 13.2 the x-distances increase by one with each step, except when descents occur. Therefore, if $T_j=1$ and the last descent occurred in C^1, we have $\delta(x^1_j)=\delta(x^2_{m+2-j})$. If the last descent occurred in C^2, then $\delta(x^1_j)=\delta(x^2_{m+2-j})+2$. This situation continues until the next descent, so the descents must alternate between the two caterpillars. Let $j=c(2)$. If $(S_2;T_2)=(0;1)$, then we have $\delta(x^1_j)=2=\delta(x^2_{m+2-j})$, since $x^1_j\sim x^1_0$ and $x^2_{m+2-j}\sim x^2_m$. If $(S_2;T_2)=(1;0)$, then $\delta(x^1_j)=3=\delta(x^2_{m+2-j})+2$. So the first descent occurs in C^1 or C^2 as described. The comment made for $S_2=1$ also shows that $t\geq 4$ is needed even for the first descent in C^2.

The necessity of the other conditions was established in Lemma 13.3, except for the condition that $t=4$ when the descent is due to arrive in C^2. The reader can refer again to Fig. 13.3. Suppose a descent occurs to r at k in C^2. In the proof of Lemma 13.3, we saw that this requires $\delta(x^1_k)=r+2$ and $\delta(x^1_{b(k)})=r+1$. By the parity condition, no descent occurs at $b(k)$, so $\delta(x^1_{b^2(k)})=r$. If $t=2$ and a descent

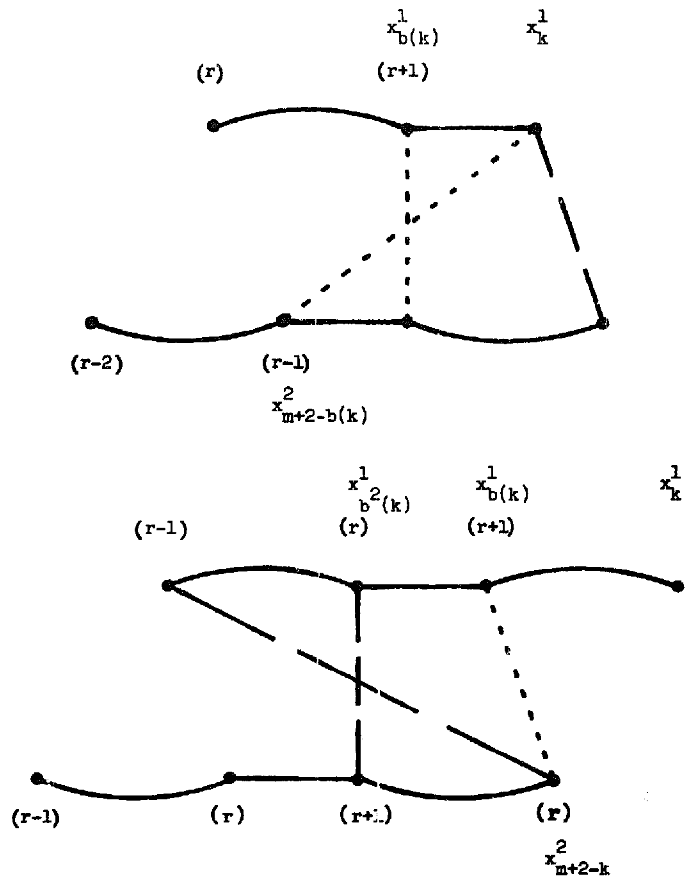

Fig. 13.4. Guaranteeing descents in C^1 and C^2.

occurs at $b^2(k)$ in C^1, then $\delta(x^2_{m+2-b^2(k)}) = r-2$. This conflicts with $\delta(x^2_{m+2-b(k)}) = r+1$.

The sufficiency is easy to establish. The two cases are illustrated in Fig. 13.4. Suppose the previous descent was in C^2 and suppose $b(k)=k-1$. We have $\delta(x^2_{m+2-b(k)}) = \delta(x^1_{b(k)})-2$. If the parity condition is also satisfied, Lemma 11.1 gives $f(x^1_k) = x^2_{m+2-b(k)}$. Hence $\delta(x^1_k) \leq \delta(x^1_{b(k)})-1$, and a descent occurs at k in C^1. On the other hand, suppose the previous descent was in C^1, $b^2(k)=b(k)-1$, and $(S_{\beta(k)}, \ldots, S_{k-1})$ has even weight at least 4. Now $\delta(x^2_{m+2-b^i(k)}) = \delta(x^2_{b^i(k)}) = r+2-i$ for $i=1,2,3$. Also $x^2_{m+2-k} = f(x^1_{b(k)}) = l(x^1_{b^3(k)})$ by Lemmas 11.1 and 7.1. Hence $\delta(x^2_{m+2-b(k)}) = r+1$ and x^2_{m+2-k} neighbors an $(r-1)$-vertex, so a descent occurs to r at k in C^2. □

The analysis of descents enables us to discuss the distribution of vertices by x-distance. We call vertices *high-valent* if they have degree at least 4. These are the vertices on the distinguished paths.

Corollary 13.5. *In an irreducible NOHO-graph $G(S;T)$, descents occur only to odd values of r. If a descent occurs to r, then G has three high-valent r-vertices and* *three high-valent* $(r+1)$*-vertices. Otherwise, G has at most two high-valent r- or* $(r+1)$*-vertices.*

Proof. This follows from the spacing of descents derived in Lemma 13.4. □

Another structural fact can be read from Fig. 13.4 if a few more edges are filled in, as in Fig. 13.5. Let a high-valent r-vertex be of *Type* i if it neighbors i high-valent $(r-1)$-vertices, and *Type* i^* if it neighbors i high-valent $(r+1)$-vertices.

Corollary 13.6. *If an irreducible* $G(S; T)$ *has two high-valent r-vertices, one is Type* 1 *and one is Type* 2. *If it has a descent to r, then its three high-valent r-vertices are one each of Types* 1^*, 2^*, *and* 3^*. *Its three high-valent* $(r+1)$*-vertices are one each of Types* 1, 2, *and* 3. *The adjacencies are as indicated in Fig.* 13.5.

This corollary enables us to distinguish between the high-valent r-vertices by their adjacencies.

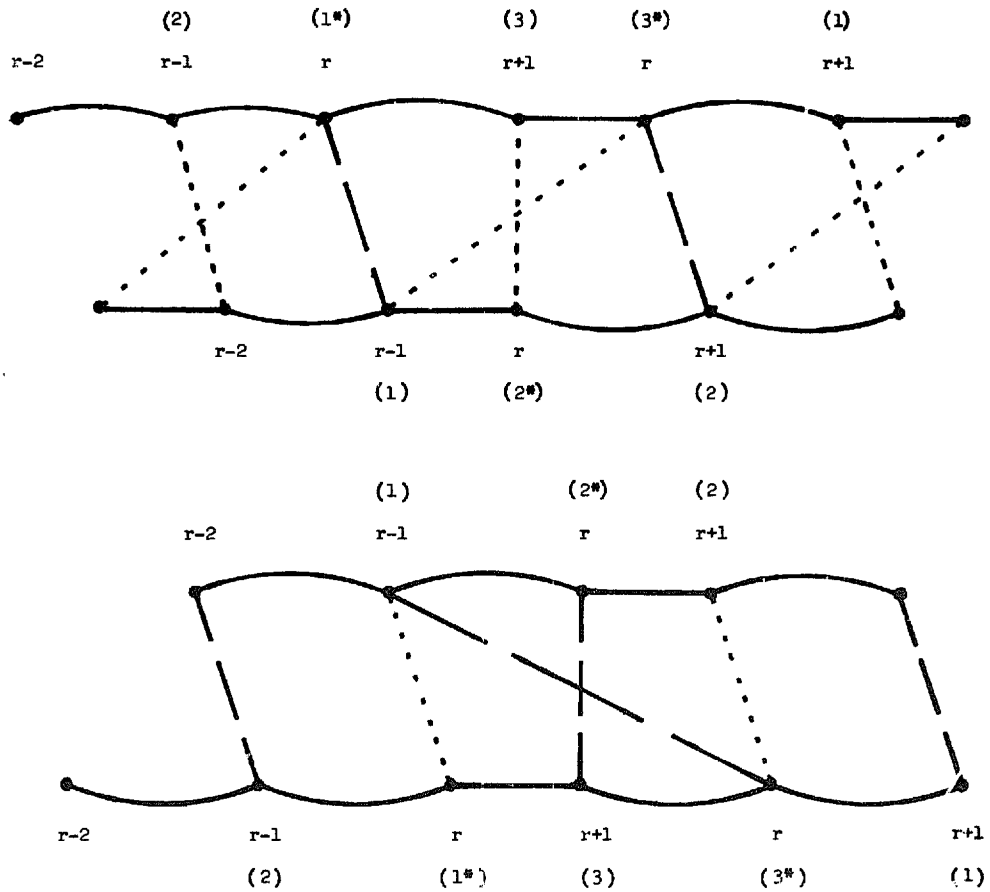

Fig. 13.5. x-distances of vertices, with types circled.

14. Conditions for isomorphism

In this section, we apply the results about twists and descents among high-valent r-vertices to determine when realizable quadruples determine isomorphic NOHO-graphs.

Theorem 14.1. *Irreducible NOHO-graphs $G(S; T)$ and $G(S'; T')$ are isomorphic if and only if $S_k = S'_k$ for $2<k<m$ or $S_k = S'_{m+2-k}$ for $2<k<m$.*

Proof. Sufficiency has been shown. Now assume the graphs are isomorphic. By reflecting and/or reversing, we can assume that the isomorphism takes y_0^1 to x_0^1 and that $S_2 = S'_2 = 1$. For each pair (r, k) the number of k-valent vertices with x-distance r is invariant under the isomorphism, as are their types. So, it suffices to show that these relationships uniquely determine $(S; T)$.

G has two vertices with x-distance 1. One of them is 3-valent; label it x_1^1. The other lies in C^2. Its label, or equivalently the index of the next 1 in $(S; T)$ after S_2, is determined by its degree, by using Remark 12.1. x_2^1 is the high-valent vertex of x-distance 2 which neighbors *both* of the vertices with x-distance 1.

We proceed similarly. Let the 'leading' caterpillar be C^i if the last descent occurred in C^i. Initially, the leading caterpillar is C^2, because we have chosen to set $S_2 = 1$. The algorithm for labelling the vertices can be read from Fig. 13.5; Corollary 13.6 insures that it works. We begin by creating the distinguished paths—equivalently, by distributing the high-valent vertices among the two caterpillars.

Suppose we have labelled all the high-valent vertices that have x-distance less than r. If there are two high-valent r-vertices, let the one of Type 2 be the next high-valent vertex in the non-leading caterpillar, and let the one of Type 1 be the next high-valent vertex in the leading caterpillar.

Suppose there are three high-valent r-vertices, where r is odd. Then there is a descent to r, and there are also three high-valent $(r+1)$-vertices. Let the leading caterpillar acquire the r-vertex of Type 2* and then the $(r+1)$-vertex of Type 2. Let the non-leading caterpillar acquire, in order, the r-vertex of Type 1*, the $(r+1)$-vertex of Type 3, the r-vertex of Type 3*, and the $(r+1)$-vertex of Type 1. Finally, let the non-leading caterpillar become the leading caterpillar.

To label the high-valent vertices, proceed as follows. Let the label of a high-valent vertex of degree d placed in C^1 be x_j^1, where $j = i + d - 3$ and x_i^1 is the label of the vertex it was assigned to follow. For those placed in C^2, let the label be x_j^2, where d is its degree, $j = i - d + 3$, and x_i^2 is the label of the vertex it follows. The labels of the high-valent vertices determine the indices of the 1's in $(S; T)$, and hence all of $(S; T)$. □

Finally, we can characterize isomorphic NOHO-graphs.

Theorem 14.2. *The quadruples determined by* $(S; T)$ *and* $(S'; T')$ *are equivalent if and only if one can be obtained from the other by a sequence of twists next to components of length* 1, *and possibly one overall reversal or twist.*

Proof. Let $\{x_j^i\}$ and $\{y_j^i\}$ be the vertices of $G = G(S; T)$ and $G' = G(S'; T')$. Sufficiency was shown in Theorem 12.4, necessity will rest on Theorem 14.1. The proof is by induction on the number of irreducible components in the two graphs. If both are irreducible, Theorem 14.1 suffices. If not, we can assume the isomorphism π takes x_0^1 to y_0^1 by initially reversing one of the pairs of sequences if necessary. Also, $(S; T)$ can be assumed to be reducible, with k the last index in some realizable segment, where $2 \leq k < m$.

We need to undo the concatenation. In defining concatenation, the first and last neighbors of the deleted 2-valent vertices were paired up and identified to unite the components. These identified vertices are now $\{x^1_{b(k+1)}, x^2_{m+2-c(k)}\}$. Deleting these separates G into two components, with vertex sets V_1 and V_2. The NOHO-graphs concatenated to get G can be obtained by letting G_i be the subgraph induced by $V_i \cup \{x^1_{b(k+1)}, x^2_{m+2-c(k)}\}$ and adding to it a vertex joined to $x^1_{b(k+1)}$ and $x^2_{m+2-c(k)}$.

Now consider the images $u_1 = \pi(x^1_{b(k+1)})$ and $u_2 = \pi(x^2_{m+2-c(k)})$ in G'. Since G and G' are isomorphic, their deletion separates G' into subgraphs V_1' and V_2' isomorphic to V_1 and V_2. The graphs G_i' produced by adding a vertex neighboring u_1 and u_2 to the subgraph induced by $V_i' \cup \{u_1, u_2\}$ are isomorphic to the NOHO graphs concatenated to make G. If we can show $\{u_1, u_2\} = \{y^1_{b(k+1)}, y^2_{m+2-c(k)}\}$, then splitting $(S'; T')$ between k and $k+1$ will produce segments realized by G_i'. That is, the labeling will be correct so that putting G_i' together to get G' is the same as concatenation, possibly with an overall twist first. The corresponding summand graphs are isomorphic; hence, by induction, the two portions of $(S'; T')$ can be transformed into the two portions of $(S; T)$ by twist, with *no reversal* because that was done initially if necessary. Therefore $(S; T)$ and (S', T') are related as desired.

It remains to be shown that $\{u_1, u_2\} = \{y^1_{b(k+1)}, y^2_{m+2-c(k)}\}$. Note that the functions b and c here are applied to indices in $(S'; T')$. Consider what pairs of vertices can disconnect G'. Removing two vertices from a single C^i does not disconnect a NOHO-graph, since $\bar{C}^j \cup F(G)$ or $\bar{C}^j \cup L(G)$ is a spanning tree, so $\{u_1, u_2\} = \{y_s^1, y^2_{m+2-t}\}$ for some s and t. All y_i^1 with $i < s$ remain connected to y_0^1 along C^1 and belong to V_1', similarly all y^2_{m+2-j} with $j > t$ belong to V_2'. There are s of the former and $m+2-t$ of the latter. However, no last-neighbor of a vertex in $C^1 \cap V_1'$ can lie in V_2', so $s \leq m+1-(m+2-t) = t-1$.

Now consider the distribution of the remaining vertices. The disconnecting vertices are high-valent, being the images of high-valent vertices. (Also, they otherwise could not disconnect the graph.) So $S_s' = T_t' = 1$. Let S_q' be the next 1 in S, and let T_r' be the preceding 1 in T'. All y_i^1 with $i \geq q$ are still joined to y_0^2 along C^1, so they belong to V_2'. All y^2_{m+2-j} with $j \leq r$ are still joined to y_0^1 along C^2, so

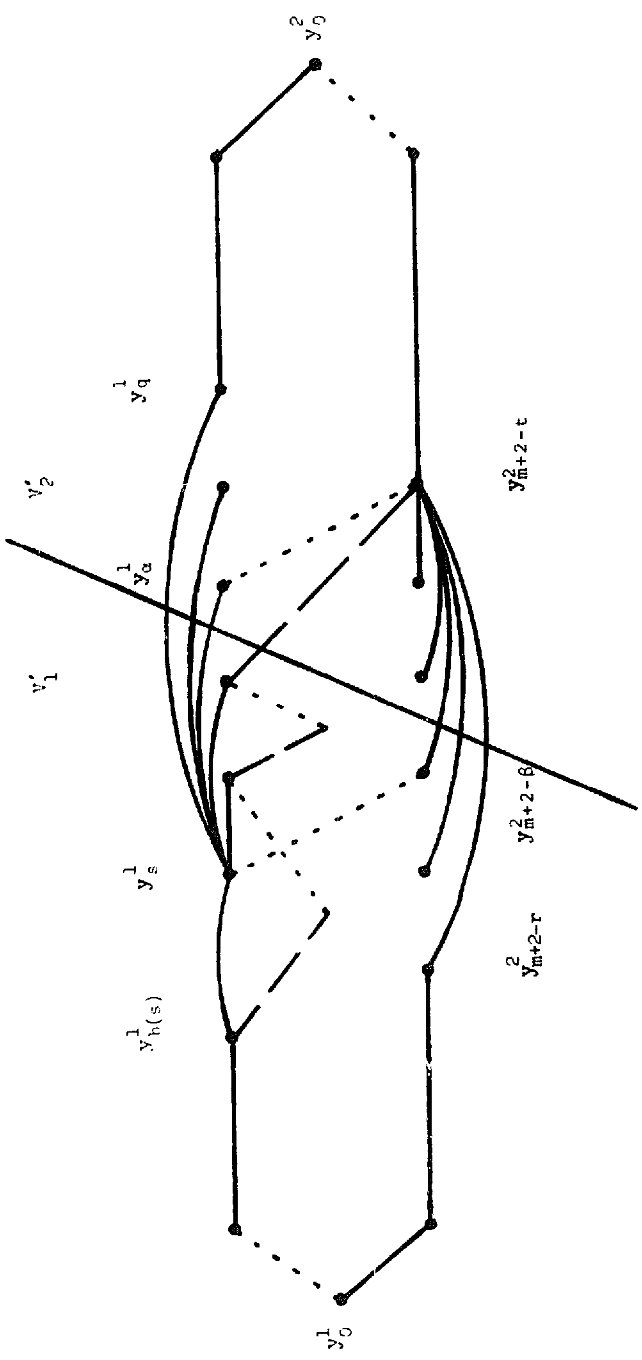

Fig. 14.1 Splitting G' into V_1' and V_1'.

they belong to V_1'. Remaining are $\{y_i^1: s<i<q\}$ and $\{y_{m+1-j}^2: r<j<t\}$, for which $S_i' = T_i' = 0$. The situation is pictured in Fig. 14.1.

Define α and β by $y_\alpha^1 = f(y_{m+2-t}^2)$, $y_{m+2-\beta}^2 = f(y_s^1)$. We claim $s<\alpha\leq q$, $r\leq\beta<t$. Furthermore, y_α^1 and $y_{m+2-\beta}^2$ mark the dividing points between V_1' and V_1', in that the vertices of V_1' consist of all $y_i^1 \neq y_s^1$ with $i<\alpha$ and all y_{m+2-j}^2 with $j\leq\beta$. To prove this, consider growing the alternating path of first and last edges starting at $y_{b(s)}^1$. Since $S_s=1$, the path U that passes through y_{s+1}^1 does not hit y_s^1. However, $S_{s+1}=\cdots=S_{q-1}=0$, so Lemma 7.1 implies that U hits every vertex of C^1 from y_{s+1}^1 through y_q^1. Since $y_{b(s)}^1 \in V_1'$ and $y_q^1 \in V_2'$, U will join the two components unless it hits y_{m+1-t}^2 and does so before hitting y_q^1. 'Before y_q^1' requires $\alpha\leq q$, since U reaches y_α^1 immediately after y_{m+2-t}^2. Continuing from y_α^1, U hits every vertex of C^1 till y_q^1. The last neighbors of these vertices lie to the right of y_{m+2-t}^2 in C^2, hence they are in V_2'. So, the y_i^1 are distributed as claimed. The symmetric argument distributes the vertices y_{m+2-j}^2 as claimed.

To complete the proof, we obtain a pair of simultaneous equations for α and β. Counting from 0 to $\alpha-1$ with s and from 2 to β, V_1' consists of $\alpha-1$ vertices of $\bar{C}^1$ and $\beta-1$ vertices of C^2. Since they can have no neighbors in V_2', these vertices are matched to each other by last edges, except that one of them matches to x_{m+2-t}^2. Hence $\alpha-2=\beta-1$. On the other hand, the total number of vertices in V_1' equals that in V_1. Since V_1 is obtained by deleting three vertices from $G(S_2,\ldots,S_k;T_2,\ldots,T_k)$, this number is $2k-1$. Together, $\alpha-\beta=1$ and $\alpha+\beta=2k+1$ give $\alpha=k+1$, $\beta=k$. So, $s=b(k+1)$ and $t=c(k)$, as desired. □

15. Enumeration of NOHO-graphs

Knowing which quadruples arise from isomorphic NOHO-graphs, we can now enumerate the non-isomorphic NOHO-graphs. This takes two steps. First, we obtain a recurrence relation for the number of equivalence classes of isomorphic NOHO-graphs under twisting. Henceforth, we refer to these simply as equivalence classes. Later, we eliminate the overcounting that remains due to the operation of reversing. Even though we now are interested only in pairs $(S;T)$, we will continue to describe them as quadruples.

Theorem 15.1. *Let u_m be the number of equivalence classes of realizable quadruples $(S_2,\ldots,S_m;T_2,\ldots,T_m)$. Then u_m satisfies the recurrence $u_m = 3u_{m-1}-u_{m-3}$.*

Proof. To facilitate the discussion of recurrence, we call a pair $(S_2,\ldots,S_m;T_2,\ldots,T_m)$ an '$(S;T)$ on m'. Also, an irreducible component of length at least two is called a *non-trivial* component. Since we are discussing $(S;T)$, it will be more natural to call the components *segments*. By Theorem 12.4 the twistable portions of $(S;T)$ are those between strings of trivial segments.

Look at the last non-trivial segment, if one exists. Note that this need not be all of the last twistable portion. There are three local changes we can make to increase the length of $(S; T)$ by one. These are:

(a) Insert $(0; 0)$ between the last two positions of the last irreducible segment.

(b) Insert $(1; 1)$ between the last two positions of the last irreducible segment and change its last position from $(1; 0)$ to $(0; 1)$ or vice versa to conform to the odd weight requirement for realizability (Lemma 11.5).

(c) Insert $(1; 1)$ *after* the last irreducible segment, leaving it alone, but lengthening the trailing string of ones.

Applying these operations to all $(S'; T')$ on $m-1$, we get $3u_{m-1}-2$ candidates for $(S; T)$ on m, since (a) and (b) cannot be applied when all the irreducible components are trivial. (The operations are equivalent to the procedure used in Corollary 11.8 to count realizable quadruples.)

Consider a realizable $(S; T)$ on m. How many times does it arise from performing operations (a), (b), (c) to a realizable $(S'; T')$ on $m-1$? Every $(S; T)$ on m arises from using (c), except those whose last non-trivial segment ends at m. If the last non-trivial segment has length two, (c) is the only way it can arise. If its length exceeds two, then $(S; T)$ also arises once from (a) or (b), regardless of whether it ends at m or earlier. To correct the count, we must (1) add one for every $(S; T)$ whose last non-trivial segment ends at m and has length two, and (2) subtract one for every $(S; T)$ whose last non-trivial segment has length more than two and ends before m.

(1) If we delete an irreducible segment of length two from positions $m-1$ and m, we get a realizable $(S'; T')$ on $m-2$. If $(S_{m-2}; T_{m-2})$ ends a non-trivial segment, then there are two choices for how the last segment could have been attached.

However, if $(S_{m-2}; T_{m-2})$ is a trivial segment, then the last segment can be twisted, meaning there is only one way to add it. Therefore, $2u_{m-2}$ counts this last collection twice, and the appropriate adjustment here is $2u_{m-2}-u_{m-3}$.

(2) Every $(S; T)$ whose last non-trivial segment ends before m and has length more than 2 arises exactly once by applying operation (c) *and* operation (a) or (b) to $(S'; T')$ on $m-2$. Every such application gives such an $(S; T)$, so the adjustment here is $2u_{m-2}-2$ (again the $(S'; T')$ of all trivial components cannot be used).

Combining the initial computation and the adjustments, we have

$$\begin{aligned} u_m &= 3u_{m-1}-2+2u_{m-2}-u_{m-3}-(2u_{m-2}-2) \\ &= 3u_{m-1}-u_{m-3}. \qquad \square \end{aligned}$$

Corollary 15.2. *Let $g(x)$ be the generating function for the number of equivalence classes of quadruples on m under twisting. Then*

$$g(x)=\frac{1-2x-2x^2}{1-3x+x^3}.$$

For the generating function, we count $u_1 = 1$ for the 4-cycle and $u_0 = 1$ for the recurrence to work. It is also possible to derive the generating function without the recurrence; in fact, it was found this way originally. This involves filling compositions with twistable portions, as in the arguments in Sections 7 and 8. Numerous summations and generating functions in two variables appear, and along the way complex analysis is used to pull a generating function in one variable from a generating function in two variables. The combinatorial argument for the recurrence is considerably simpler.

We still need to eliminate some overcounting. To do this we consider symmetric and reversible pairs $(S; T)$. Recall that reversible $(S; T)$ have been characterized as those for which $S_k = S_{m+2-k}$ and $T_k = T_{m+2-k}$. Symmetric $(S; T)$ are those for which $S_k = T_{m+2-k}$.

A NOHO-graph which becomes neither reversible nor symmetric under any sequence of legal twists is counted twice by u_m, since reversing takes it to a different equivalence class. Before dividing by two, we must add the number of classes containing symmetric or reversible quadruples. Keep in mind that each such class represents only one underlying graph, no matter how many symmetric or irreducible quadruples it contains.

Lemma 15.3. *Let q_m be the number of NOHO-graphs on $2m+2$ vertices that can be represented by both symmetric and reversible quadruples. Then $q_m = u_{\lfloor m/2 \rfloor}$.*

Proof. Symmetry and reversibility each require symmetry of the composition representing the lengths of irreducible parts. Together they also place a restriction on what can happen in the middle of $(S; T)$. Whether m is even or odd, the combination of reversibility and symmetry prevents a non-trivial quadruple from straddling the center or two of them from meeting at the center. In the latter case, recall that no twist can occur between those parts. If such a quadruple is reversible, the opposite ends of the central irreducible part or central pair of parts must be $(1; 0)$ and $(1; 0)$. For a symmetric quadruple, they must be $(1; 0)$ and $(0; 1)$. They can't have it both ways.

This means a symmetric reversible quadruple must have one (m even) or two (m odd) trivial parts at the center to allow the needed twists (see Fig. 15.1). It then suffices to place a representative of any equivalence class in the first half and reflect it into the second half. A twist in the middle will make that a reversible quadruple.

When m is even, the $(1; 1)$ in the middle occupies position $\frac{1}{2}m + 1$. It serves as position $m' + 1$ for the segment in the left half of $(S; T)$, so $m' = \frac{1}{2}m$. When m is odd, the $(1, 1; 1, 1)$ in the middle occupies positions $\frac{1}{2}(m+1)$ and $\frac{1}{2}(m+3)$. This time $\frac{1}{2}(m+1)$ serves as $m'+1$, so $m' = \frac{1}{2}(m-1)$. In either case, the number of symmetric reversible quadruples is $u_{\lfloor m/2 \rfloor}$. □

Lemma 15.4. *Let r_m be the number of NOHO-graphs on $2m+2$ vertices represent-*

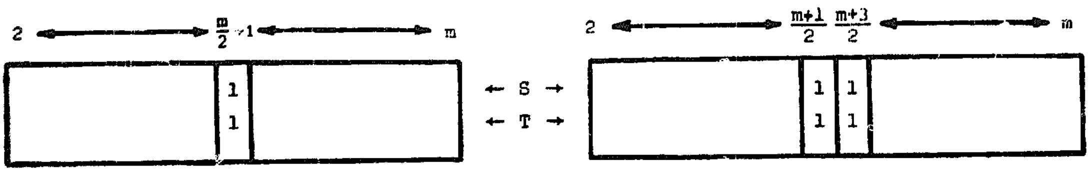

Fig. 15.1. Parity considerations.

able by a reversible quadruple but no symmetric quadruples. Then $r_{2k+1} = u_{k+1} - u_k$ *and* $r_{2k} = r_{2k+1}$.

Proof. Such a reversible quadruple cannot have any trivial parts in the middle, since then a twist around them would make the quadruple symmetric. Therefore, when m is odd two non-trivial irreducible parts meet in the middle, and when m is even an irreducible part straddles the middle. Suppose $m = 2k+1$, so the central positions are $k+1$ and $k+2$. Again we fill the first half. This time we have freedom through $k+1$, so we pretend $k+2$ gets (1; 1) and count the number u_{k+1} of classes on $k+1$. However, we must discard those which end with an extra trivial part (1; 1) at $k+1$. There are u_k of those.

When $m = 2k$, there is a central position. As noted above, it mustn't contain a trivial part, so some non-trivial irreducible part straddles the center. Outside the central position S and T in the central irreducible part have even weight, because the same thing happens on both sides. So, to be realizable, the center position gets (1; 1). Take any reversible, non-symmetric quadruple on $2k+1$. It has two parts meeting as (00; 11) or (11; 00) in the center. The weight of S and T in each irreducible part is odd; uniting both parts yields S and T with even weight. Replace the central two pairs by (1; 1). This changes the parity in S and T to yield one central irreducible part, counted by r_{2k}. This process is clearly reversible. The central (1; 1) turns into whichever is the unique choice of (00; 11) and (11; 00) to make S; T realizable (and reversible). Therefore, $r_{2k} = r_{2k+1}$. This is summarized in Fig. 15.2. □

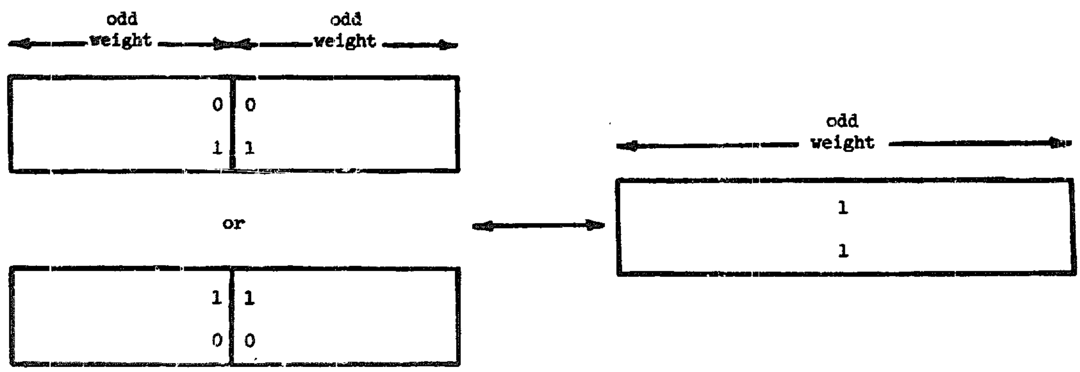

Fig. 15.2. Generating reversible quadruples.

Lemma 15.5. *Let s_m be the number of NOHO-graphs on $2m+2$ vertices representable by a symmetric quadruple but no reversible quadruples. Then $s_{2k}=r_{2k}$, and $s_{2k-1}=r_{2k}+r_{2k-1}=u_{k+1}-u_{k-1}$.*

Proof. As in Lemma 15.4, such a quadruple has no trivial parts in the center. When $m=2k$, there is an odd length symmetric quadruple in the center. These quadruples can be obtained from those counted by r_{2k} by replacing the central $(1;1)$ by $(0;0)$ and performing a (normally illegal) twist at the center. The change to $(0;0)$ is appropriate because the movement in the twisted end-position of the central quadruple changes the parity. Obviously, the quadruple is now symmetric, and this is a 1–1 correspondence.

If $m=2k-1$, the quadruple can have a central symmetric irreducible part of even length, or have two non-trivial irreducible parts meet there. In the former case, flipping and inserting a central $(1;1)$ gives a 1–1 correspondence with non-symmetric reversible quadruples on $2k$, yielding the r_{2k} term. In the latter case, only the flip around the center is needed to switch between non-symmetric reversible and non-reversible symmetric quadruples on $2k-1$. □

Theorem 15.6. *Let t_m be the number of non-isomorphic NOHO-graphs on $2m+2$ vertices. Then*

$$t_m=\tfrac{1}{2}(u_m+u_{\lceil m/2\rceil+1}+u_{\lfloor m/2\rfloor+1}-u_{\lfloor m/2\rfloor})$$

Proof. As noted in the paragraph before Lemma 15.3, $t_m=\frac{1}{2}(u_m+r_m+s_m+q_m)$. By Lemmas 15.3–15.5, this yields

$$t_{2k}=\tfrac{1}{2}(u_{2k}+u_{k+1}-u_k+u_{k+1}-u_k+u_k)=\tfrac{1}{2}(u_{2k}+2u_{k+1}-u_k),$$

$$t_{2k-1}=\tfrac{1}{2}(u_{2k-1}+u_k-u_{k-1}+u_{k+1}-u_{k-1}+u_{k-1})=\tfrac{1}{2}(u_{2k-1}+u_{k+1}+u_k-u_{k-1}).\quad \square$$

Corollary 15.7. *Let $h(x)$ be the generating function $\sum t_m x^m$. Then*

$$h(x)=\tfrac{1}{2}g(x)+\frac{1}{2x^3}[g(x^2)(1+x)(1+x-x^3)-1-2x-2x^2].$$

Proof. Follows directly from Theorem 15.6 by standard techniques for manipulating generating functions. □

Finally we summarize these results by listing values for the numbers of equivalence classes, symmetric reversible NOHO-graphs, reversible non-symmetric NOHO-graphs, symmetric non-reversible NOHO-graphs, and total number of NOHO-graphs on $2m+2$ vertices.

Table 1. Enumerating NOHO-graphs on $2m+2$ vertices

m	u_m	q_m	r_m	s_m	t_m
2	1	1	0	0	1
3	2	1	0	0	2
4	5	1	1	1	4
5	14	1	1	4	10
6	40	2	3	3	24
7	115	2	3	12	66
8	331	5	9	9	177
9	953	5	9	35	501
10	2744	14	26	26	1405
11	7901	14	26	101	4021
12	22750	40	75	75	11470
13	65506	40	75	291	32956

16. Related gossip questions

Golumbic [2] and Harary and Schwenk [3] noted that any connected multigraph with n vertices, $2n-4$ edges, and a 4-cycle admits an edge-ordering which pools all the information. The question remained open for many years whether every solution of the gossip problem using $2n-4$ edges contains a 4-cycle. Bumby [1] settled this in the affirmative. NOHO-graphs contain many 4-cycles.

Corollary 16.1. *Every edge of a NOHO-graph appears in at least two 4-cycles.*

Proof. In Corollary 7.2, we saw that the first and last neighbors of any vertex in a NOHO-graph (with edge-ordering) are joined by a path of two edges in the opposite extended caterpillar. Including the first and last edges to that vertex yields a 4-cycle. A first or last edge appears in the 4-cycle generated by each of its end-points. By Lemma 7.1 an edge in one of the caterpillars appears in the 4-cycles generated by the last neighbor of its lesser endpoint and the first neighbor of its greater endpoint (lesser or greater in the canonical numbering). □

A NOHO-graph also has cycles containing one edge from each caterpillar. For example, when $P_i < P_j$ with $i<j$, $(x_i^1, x_j^1, f(x_j^1), f(x_i^1))$ is a 4-cycle, by Corollary 6.2, and there are others. It would be interesting to determine what other properties of NOHO-graphs must be added to get a graphical characterization of NOHO-graphs independent of the edge-orderings. Planar, Hamiltonian, and bipartite are good properties to start with. In fact, because incident first-last edges do determine 4-cycles in the other caterpillar, the planar representation constructed in Section 5 of Part I is one in which all the faces are 4-cycles, so we can require that also. Unfortunately, not even adding the requirement that the graph

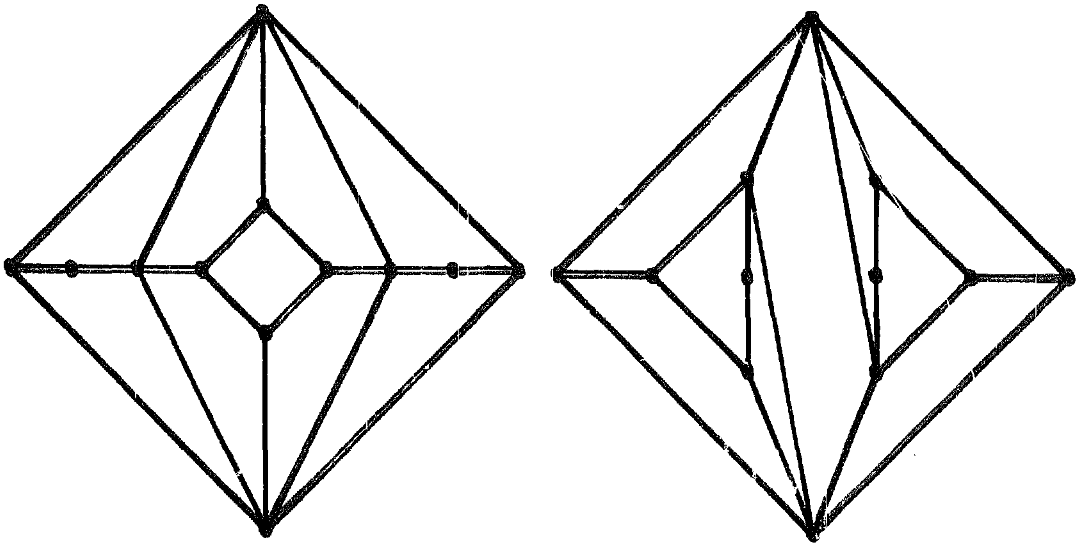

Fig. 16.1. NOHO-graphs?

have two 2-valent vertices is sufficient. For example, the two plane graphs in Fig. 16.1 are Hamiltonian and bipartite, with all their faces being 4-cycles. They also have the same degree sequences. However, only the graph on the right is a NOHO-graph.

Beyond NOHO, we can require that no transmission of information be duplicated. In other words, there must be exactly one increasing path from each vertex to every other vertex. Lenstra [4] has considered this question and shown that such schemes exist when 4 divides n. He uses an inductive construction that takes $\frac{1}{2}n \log n + O(1)$ calls. In a forthcoming note [7], we present a scheme that uses only $\frac{9}{4}n - 6$ calls.

Paradoxically, forbidding wastage requires more calls, if indeed the problem can be solved at all. This follows from the next result. For undirected edges, 'NODUP' implies NOHO, since any increasing cycle would contain two increasing paths from the bad vertex to its last neighbor on that cycle. So, except for the special graphs when $n = 4$ or $n = 8$, 'NODUP' requires more than $2n - 4$ calls, since

Corollary 16.2. *Every NOHO-graph duplicates some transmission(s).*

Proof. Consider any NOHO-graph with edge-ordering. We claim there are two increasing paths from x_2^1 to $l(x_1^1)$. Certainly $(x_2^1, x_1^1, l(x_1^1))$ is such a path; we claim $f(x_2^1) \sim l(x_1^1)$, so that $(x_2^1, f(x_2^1), l(x_1^1))$ is another such path. As usual, there are two cases. The value of T_2 determines whether x_m^2 lies on the distinguished path in $\bar{C}^2$. If $T_2 = 1$, then $x_m^2 = l(x_0^1)$, while if $T_2 = 0$, then the earlier x_j^2 to which x_m is joined is $l(x_0^1)$. We use Corollary 7.2 again. If $T_2 = 1$, it yields $l(x_1^1) = x_{m+2-c(2)}^2$ and $f(x_2^1) = x_m^2$. If $T_2 = 0$, it yields $l(x_1^1) = x_m^2$ and $f(x_2^1) = x_{m+2-c(m)}^2$. In either case, the two paths exist as claimed. $\square$

Finally, we describe a generalization of the problem considered here. Consider an n by n 'transmission matrix' on vertices $\{v_1, \ldots, v_n\}$, with entries from $\{1, 0, -1\}$. If $a_{ij} = 1$, we require an increasing path from v_i to v_j. If $a_{ij} = -1$, we forbid such a path. If $a_{ij} = 0$, we don't care. We ask whether a calling scheme satisfying the matrix exists, what is the least number of calls in such a scheme, what schemes achieve the minimum, and so on. The original gossip problem results when diagonal entries are 0 and off-diagonal entries are 1. Changing the diagonal entries to -1 yields the NOHO-condition. The problem with ones above the diagonal and zeros on or below it is clearly optimized by an increasing path of $n-1$ edges. For a matrix in block diagonal form, we require the sum of the calls required by the smaller problems. Here's another example:

Corollary 16.3. *Consider a transmission matrix with $a_{ii} = 0$, $a_{ij} = 0$ for $i > r \geq j$, and all other $a_{ij} = 1$. The smallest graph solving this gossip problem has $2n-7$ edges. This remains true if a_{ii} are changed to -1, as long as n and r are even. If they are not, no scheme is feasible.*

Proof. Take an ordinary $(2r-4)$-edge solution H_1 on $\{v_1, \ldots, v_r\}$, and an ordinary $(2n-2r-4)$-edge solution H_2 on $\{v_{r+1}, \ldots, v_n\}$. Order the edges so all those of H_2 occur after all those of H_1. Add an edge joining a vertex of the last edge in H_1 to a vertex of the first edge in H_2, and let it occur between them. This uses $2n-7$ calls and satisfies the matrix.

To show optimality, take any solution and delay all edges not wholly within $\{v_1, \ldots, v_r\}$, in order, until after every edge within that set. The resulting scheme still satisfies the transmission matrix. But now it must consist of an ordinary scheme on r vertices, followed by at least one connecting edge and a solution on $n-r$ vertices. So, there are at least $2n-7$ calls.

If $a_{ii} = -1$, then the graphs in the H_1, H_2 construction must satisfy NOHO. This requires r and $n-r$ even. □

There are innumerable variations.

References

[1] R.T. Bumby, A problem with telephones, SIAM J. Algebraic Discrete Methods 2 (1981) 13–19.
[2] M.C. Golumbic, The general gossip problem, IBM Research Report, 1974.
[3] F. Harary and A.J. Schwenk, The communication problem on graphs and digraphs, J. Franklin Inst. 297 (1974).
[4] H.K. Lenstra, Private communication.
[5] D.B. West, On a class of solutions to the gossip problem I, Discrete Math. 39 (1982). 307–326.
[6] D.B. West, On a class of solutions to the gossip problem II. Discrete Math. 40 (1982). 87–113.
[7] D.B. West, Gossiping without duplicate transmissions, SIAM J. Algebraic Discrete Methods, to appear.