ON PRIMITIVE GRAPHS AND OPTIMAL VERTEX ASSIGNMENTS

by

R. L. Graham,
Bell Telephone Laboratories, Incorporated
Murray Hill, New Jersey

I. INTRODUCTION

The results presented here are an outgrowth of an investigation on the subject of "cube-numbering" and some of its generalizations. "Cube-numbering" refers to the general problem of assigning a given set of real values to the vertices of an n-cube (or more generally, some graph G) so that the sum over all edges e of the n-cube (or G) of a given function of the difference of the values of the two endpoints of e is minimized. This concept arises in the study of certain optimal binary codes (cf. [4], [5], [7]). In this note we extend the set of admissible graphs G (of which the n-cube is now a special case) and express the preceding question of optimal vertex assignments in number-theoretic terms. The solution to the corresponding number theory problem is obtained by studying the structure of a special class of graphs we call primitive. Although the few facts established here about primitive graphs are sufficient to completely answer our vertex assignment problem, it will be seen that many interesting open questions remain.

II. CUBE-NUMBERING

One of the first problems to be asked in this subject was the following. Let C^n denote the graph formed from the 1-skeleton of a unit n-cube. In other words, the set of vertices $V(C^n)$ of the graph consists of the set of the 2^n distinct binary n-tuples; a pair of vertices is joined by an edge if the two binary n-tuples differ in exactly one coordinate. Given a set* of real numbers $A = \{a_0 \le a_1 \le \dots \le a_{2^n-1}\}$, the problem is to determine the value of the expression

$$v(A) \equiv \min_{\varphi} \sum_{e=\{a,b\}\in E(C^n)} |\varphi(a) - \varphi(b)| \qquad (1)$$

where $E(C^n)$ denotes the set of edges of C^n and φ ranges over all 1-1 maps of $V(C^n)$ onto A.

This was first solved by L. H. Harper [2] for the case $a_i = i$. Later, K. Steiglitz and A. J. Bernstein [7] showed

*Where repetition is allowed.

that the arguments in [2] can also be applied to the case of general A. A particular map φ which achieves the minimum is given by

$$\varphi(\alpha_0,\ldots,\alpha_{n-1}) = a_\alpha \tag{2}$$

where $\alpha_i = 0$ or 1 and $\alpha = \sum_{i=0}^{n-1} \alpha_i 2^i$; we shall call this map the canonical assignment.

III. COMPLETE PAIRINGS

Let $X = \{x_0, x_1, \ldots, x_{2m-1}\}$ denote a set[†] of 2m real numbers. We define a pairing P of X to be a partition of X into two sets X_0 and X_1 of cardinality m together with a map θ of X_0 onto X_1. We define the value of P by

$$v(P) \equiv \sum_{x \in X_0} |x - \theta(x)|. \tag{3}$$

If m were even, we could now form pairings P_0 and P_1 of X_0 and X_1 respectively with values $v(P_0)$ and $v(P_1)$. Furthermore, if m/2 were even we could continue this process still another step forming pairings $P_{00}, P_{01}, P_{10}, P_{11}$ of $X_{00}, X_{01}, X_{10}, X_{11}$ respectively, where $X_0 = X_{00} \cup X_{01}$ and $X_1 = X_{10} \cup X_{11}$ are the partitions induced by the pairings P_0 and P_1. Finally, in the extreme case in which $m = 2^n$, we could continue for n steps, each time forming a pairing of all sets created by partitions of previous pairings. At the end, the original set X has been partitioned into singletons. We define this process to be a complete pairing P* of X and we define the total value of P* by

$$V(P^*) = \sum_P v(P) \tag{4}$$

where the sum is taken over all pairings P which occurred in the complete pairing P*.

As an example, for the set X = {0,1,3,6,6,7,9,11} we have the following complete pairing P*:

[†] As before, we allow repetition.

$$P: \begin{matrix} 0 & 3 & 6 & 11 \\ \updownarrow & \updownarrow & \updownarrow & \updownarrow \\ 1 & 7 & 6 & 9 \end{matrix}$$

$$v(P) = 1+4+0+2 = 7$$

$$P_0: \begin{matrix} 0 & 6 \\ \updownarrow & \updownarrow \\ 11 & 3 \end{matrix} \qquad v(P_0) = 11+3 = 14$$

$$P_1: \begin{matrix} 1 & 7 \\ \updownarrow & \updownarrow \\ 6 & 9 \end{matrix} \qquad v(P_1) = 5+2 = 7$$

$$P_{00}: \begin{matrix} 0 \\ \updownarrow \\ 6 \end{matrix} \quad v(P_{00}) = 6$$

$$P_{01}: \begin{matrix} 3 \\ \updownarrow \\ 11 \end{matrix} \quad v(P_{01}) = 8$$

$$P_{10}: \begin{matrix} 1 \\ \updownarrow \\ 7 \end{matrix} \quad v(P_{10}) = 6$$

$$P_{11}: \begin{matrix} 6 \\ \updownarrow \\ 9 \end{matrix} \quad v(P_{11}) = 3$$

$$V(P^*) = 7 + (14+7) + (6+8+6+3) = 51$$

Example 1

The general question is to determine

$$\mu(X) \equiv \min_{P^*} V(P^*) \tag{5}$$

where the minimum is taken over all complete pairings P^* of X. For example, for the set X of Example 1, it is not hard to check that $\mu(X) = 45$.

A little thought shows that for any $A = \{a_0 \leq \cdots \leq a_{2^n-1}\}$

$$\mu(A) \leq \nu(A). \tag{6}$$

One of our results (cf. Theorem 3) will be that equality always holds in (6). In order to attack this problem, we first examine a special class of graphs.

IV. PRIMITIVE GRAPHS

We consider a connected graph† G with $V(G)$ and $E(G)$ the vertex and edge sets, respectively. We begin by giving several definitions.

Definition: A subset $C \subseteq E(G)$ is said to be a cutset of G if:

(i) The graph with vertex set $V(G)$ and edge set $E(G) - C$ is disconnected.

(ii) (i) does not hold if C is replaced by any proper subset of C.

†For the more standard concepts in graph theory, the reader is referred to [6].

We say that C is a simple cutset if no two edges of C have a common vertex. We call G decomposable if G has a simple cutset; otherwise we say that G is indecomposable. G will be called primitive if G is indecomposable but every proper subgraph of G is decomposable. Finally we say that G is completely decomposable if every subgraph of G (including G itself) is decomposable.

We give several examples to illustrate these concepts.

Example 2: $G = K_3$, the complete graph on 3 points (cf. Fig. 1).

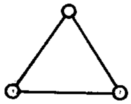

Figure 1

G is certainly indecomposable. Also, if any edge of G is removed then the remaining graph is completely decomposable. Hence G is primitive. Notice that any graph (except K_3) which contains K_3 as a subgraph can be neither primitive nor completely decomposable.

Example 3: $G = C^n$, the 1-skeleton of the n-cube.

If we take C to be the set of all edges of G parallel to a fixed edge, then C is a cutset of G whose removal leaves a graph consisting of two disjoint copies of C^{n-1}. Hence, since C^1 is trivially completely decomposable then by induction $G = C^n$ is completely decomposable.

We point out an important, though obvious, fact.

Fact 1: If G is not completely decomposable then G contains a primitive subgraph.

In particular, any indecomposable graph G must contain a subgraph which is primitive (possibly G itself). However, knowing that a primitive subgraph exists and exhibiting it may be two different things. To illustrate this point, consider the following:

Exercise: Let G be defined by Fig. 2.

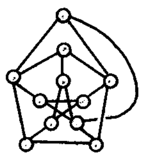

Figure 2

(i) Show G is indecomposable.

(ii) Show G is not primitive.

(iii) Find a primitive subgraph of G.

This graph was first suggested by D. Kleitman as a candidate for a primitive graph. Any solution in less than a half hour is to be considered excellent.

Definition: Let e = {x,y} be an edge of a primitive graph G and let G′ denote the graph with V(G′) = V(G), E(G′) = E(G) - {e}. We say that e is a regular edge of G if G′ has both a simple cutset no edge of which contains x and also a simple cutset no edge of which contains y.

We next show how two primitive graphs may be combined to yield a new primitive graph. Let G be a primitive graph with a regular edge e and let H be a primitive graph with a vertex z of degree 2 (cf. Fig. 3).

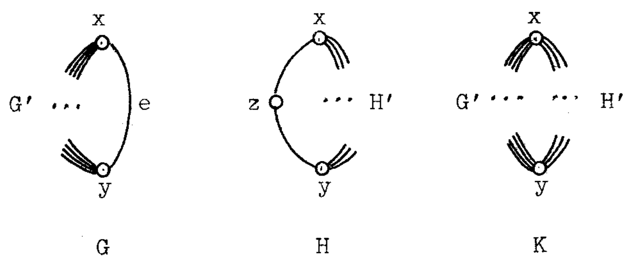

Figure 3

We assume that V(G) ∩ V(H) = {x,y} and the two edges of E(H) incident to z are {z,x} and {z,y}. We form the graphs G′ and H′ as follows:

$$V(G') = V(G), \quad E(G') = E(G) - \{e\},$$

$$V(H') = V(H) - \{z\}, \quad E(H') = E(H) - \{z,x\} - \{z,y\}.$$

Finally, we form the graph K = G′ ∪ H′ by

$$V(K) = V(G') \cup V(H'),$$

$$E(K) = E(G') \cup E(H').$$

Theorem 1: K is primitive.

Proof: First, assume K is decomposable. Let D be a simple cutset for K and define K′ by V(K′) = V(K), E(K′) = E(K) - D. If x and y are in different components of K′ then D ∩ E(H) must be a simple cutset for H′. But since H is primitive then any simple cutset for H′ must contain both x and y since otherwise by choosing either {z,x} or {z,y} we would have a simple cutset for H which is impossible. Therefore D contains both x and y. Hence, we can construct a simple cutset for G substituting e for $D \cap H$ in D. This is a contradiction and we conclude that <u>K must be indecomposable</u>.

Next, let $f \in E(K)$ and define K'' by $V(K'') = V(K)$, $E(K'') = E(K) - \{f\}$. There are two cases:

(i) <u>$f \in E(G')$</u>. Define G'' by $V(G'') = V(G)$, $E(G'') = E(G) - \{f\}$. Let D be a simple cutset for G''. If $e \notin D$ then D is also a simple cutset for K''. Furthermore, we can continue to delete simple cutsets from $E(G'') \subset E(K'')$ until all that remains in $E(K'')$ is just $E(H')$. We can now complete the decomposition since H' is completely decomposable and therefore K'' is completely decomposable. If $e \in D$ then no edge of $D' = D - \{e\}$ contains either x or y. Let D'' be a simple cutset for H'. Thus, $D' \cup D''$ is a simple cutset for K'' and, as before, it is not difficult to see that K'' is also completely decomposable in this case.

(ii) <u>$f \in E(H')$</u>. Define the graph H'' by $V(H'') = V(H)$, $E(H'') = E(H) - \{f\}$. Let D_1 be a simple cutset for H''. Form the graph H_1 with $V(H_1) = V(H'')$, $E(H_1) = E(H'') - D_1$. If x and y are in the same component of H_1 then D_1 is also a simple cutset for K''. In general we can continue forming graphs H_i, $2 \leq i \leq m$, so that $V(H_i) = V(H'')$, $E(H_i) = E(H_{i-1}) - D_i$, where D_i is a simple cutset for H_{i-1}, $2 \leq i \leq m$, and for the first time, x and y are in different components of H_m. Thus, for $1 \leq i \leq m - 1$, D_i is a simple cutset for $G' \cup H_{i-1}$. Also, <u>exactly one</u> of $\{z,x\}$, $\{z,y\}$ must belong to D_m. Since e was a regular edge G then we can find a simple cutset D^* for G', all of whose edges are disjoint from those of D_m. Therefore, $D_m \cup D^*$ is a simple cutset for $G' \cup H_m$ and it is easily seen that we can now completely decompose $G' \cup H_m$. This establishes the complete decomposability of K'' in this case.

Hence, in all cases K'' is completely decomposable. Since f was arbitrary, then K is primitive and the theorem is proved.

It is not difficult to show that if all edges of G and H are regular then all edges of K are also regular. Thus, we can use Theorem 1 to generate infinite families of primitive graphs. We list some of the simpler members of several of these families in Fig. 4.

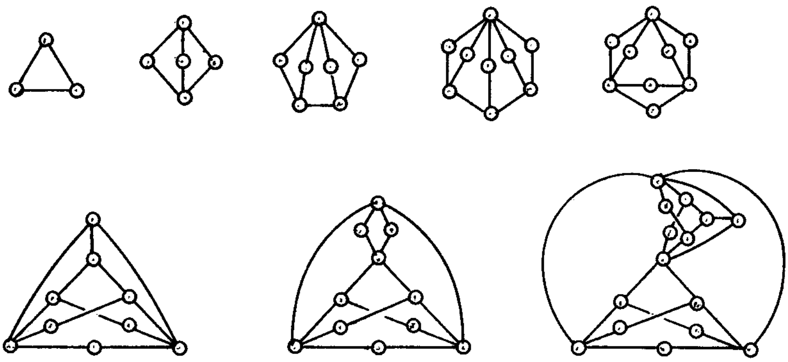

Figure 4

For a nonnegative integer k, let w(k) denote the number of 1's in the binary expansion of k. Define W(k) to be w(0) + w(1) + ... + w(k). The next result restricts the number of edges of a completely decomposable graph.

Theorem 2: Let G be a completely decomposable graph with n vertices. Then

$$|E(G)| \leq W(n-1). \tag{7}$$

This bound is best possible.

Proof: The fact that (7) is the best we could hope for follows from considering the graph G_n defined as follows:

(i) $V(G_n) = \{0,1, \ldots, n-1\}$,

(ii) $\{i,j\} \in E(G_n)$ if and only if the binary expansions of i and j differ in exactly one place.

It is easily seen that

$$|E(G_n)| = W(n-1).$$

Since G_n is a subgraph of the N-cube C^N for $2^N \geq n$ then the complete decomposability of C^N (Ex. 3) implies that G_n is also completely decomposable.

The first few values of the right-hand side of (7) are given in Table 1.

n	0	1	2	3	4	5	6	7	8
w(n)	0	1	1	2	1	2	2	3	1
W(n)	0	1	2	4	5	7	9	12	13

Table 1

The theorem holds for $n = 1$ and 2 by the definition of a graph, i.e., loops and multiple edges are prohibited. For $n = 3$, there is only one graph with 3 vertices and 3 edges (cf. Fig. 1) and since this is not (completely) decomposable then the theorem also holds in this case. Assume the theorem has been established for all graphs with $< n$ vertices and suppose there exists a completely decomposable graph G with n vertices and $e \geq W(n-1) + 1$ edges. By definition, G must have a simple cutset D. If we delete the edges of D from E(G) we must be left with exactly two components G_1, G_2, say with $|V(G_i)| = n_i$, $|E(G_i)| = e_i$, $i = 1,2$. Condition (ii) in the definition of a cutset prevents the formation of more than two components when D is removed from E(G). We can assume without loss of generality that $0 < n_1 \leq n_2$. It is easy to see that $|D| \leq \min\{n_1,n_2\} = n_1$ and hence,

$$e_1 + e_2 + n_1 \geq e \geq W(n-1) + 1. \tag{8}$$

However, if either

$$e_1 \geq W(n_1-1) + 1 \quad \text{or} \quad e_2 \geq W(n_2-1) + 1$$

then by the induction hypothesis, one of the two components is not completely decomposable and this is a contradication. Thus, we must have

$$e_1 \leq W(n_1-1), \quad e_2 \leq W(n_2-1). \tag{9}$$

Hence, combining (8) and (9) with the fact that $n_1 + n_2 = n$ we obtain

$$W(n_1-1) + W(n_2-1) + n_1 \geq W(n_1+n_2-1) + 1, \quad 0 < n_1 \leq n_2. \tag{10}$$

Therefore, the proof of the theorem will be completed by proving the following lemma which contradicts (10).

<u>Lemma 2</u>: Let r and s be integers ≥ 0. For a 1-1 map $\varphi : \{0,1,\ldots,r\} \to \{s,s+1,\ldots,s+r\}$ define $\delta(\varphi)$ by

$$\delta(\varphi) = \min_{0\leq k\leq r} \{w(\varphi(k)) - w(k)\}. \tag{11}$$

Then

(i) There exists φ such that $\delta(\varphi) \geq 0$.

(ii) If $s > r$ then there exists φ such that $\delta(\varphi) \geq 1$.

Proof: The proof proceeds by induction on r. For $r = 0$ and all s the lemma holds. Assume the lemma holds for $r - 1$ and all $s \geq 0$. Also, for $s = 0$, (i) holds with φ the identity map and (ii) holds vacuously. Hence, we may assume $s > 0$. Let p satisfy $\log_2(r+1) \leq p < 1 + \log_2(r+1)$. In the set $\{s,s+1,\ldots,s+r\}$ there is exactly one integer which is divisible by 2^p. Denote this integer by $2^q \cdot u$ where $q \geq p$ and u is odd. We can write $\{s,s+1,\ldots,s+r\}$ as $\{2^q \cdot u-r+x,\ldots,2^q \cdot u,\ldots,2^q \cdot u+x\}$ for some x, $0 \leq x \leq r$. If $x = r$ then (i) and (ii) hold by mapping $k \to 2^q \cdot u + k$ for $0 \leq k \leq r$. Hence we can assume $0 \leq x < r$.

If $s \leq r$ so that the two sets overlap then we only need to establish (i). Partition the two sets into $\{0,1,\ldots,s-1\} \cup \{s,\ldots,r\}$ and $\{s,\ldots,r\} \cup \{r+1,\ldots,s+r\}$. By the induction hypothesis there is a map $\varphi : \{0,1,\ldots,s-1\} \to \{r+1,\ldots,s+r\}$ with $\delta(\varphi) \geq 1 > 0$. We can combine this with the identity map of $\{s,\ldots,r\}$ into $\{s,\ldots,r\}$ to obtain a map $\varphi' : \{0,1,\ldots,r\} \to \{s,s+1,\ldots,s+r\}$ with $\delta(\varphi') = 0$. This establishes (i) in the case $s \leq r$.

If $s > r$ then it suffices to establish (ii) since this implies (i). Partition the two sets into $\{0,1,\ldots,x\} \cup \{x+1,\ldots,r\}$ and $\{2^q u-r+x,\ldots,2^q u-1\} \cup \{2^q \cdot u,\ldots,2^q \cdot u+x\}$. Let $\varphi_1 : \{0,1,\ldots,x\} \to \{2^q \cdot u,\ldots,2^q \cdot u+x\}$ by $\varphi_1(k) = 2^q u + k$, $0 \leq k \leq x$. Certainly $\delta(\varphi_1) \geq 1$. Since u is odd, the lemma is proved if we can find $\varphi_2 : \{x+1,\ldots,r\} \to \{2^q u-r+x,\ldots,2^q u-1\}$ with $\delta(\varphi_2) \geq 1$.

By definition $r + 1 \leq 2^p \leq 2^q$. Since u is odd we have

$$w(j) + w(2^q \cdot u-1-j) = w(u-1) + q, \quad x + 1 \leq j \leq r. \quad (12)$$

But

$$w(\varphi_2(k)) - w(k) \geq 1, \quad x + 1 \leq k \leq r,$$

$$\text{iff} \quad w(u-1) + q - w(k) - (w(u-1) + q - w(\varphi_2(k))) \geq 1, \quad x + 1 \leq k \leq r,$$

$$\text{iff} \quad w(2^q \cdot u-1-k) - w(2^q \cdot u-1-\varphi_2(k)) \geq 1, \quad x + 1 \leq k \leq r.$$

Hence, it suffices to find a map

$$\varphi_3 : \{0,1,\ldots,r-x-1\} \to \{2^q \cdot u-1-(x+1),\ldots,2^q \cdot u-1-r\}$$
$$\text{with } \delta(\varphi_3) \geq 1.$$

But $x \geq 0$ and these sets are disjoint. Hence by the induction hypothesis, such a map φ_3 exists. This completes the proof of the lemma.

We remark that the contradiction to (10) is obtained as follows. Let $\varphi : \{0,1,\ldots,n_1-1\} \to \{n_2,\ldots,n_2+n_1-1\}$ with $\delta(\varphi) \geq 1$. This is possible by Lemma 2 since $0 < n_1 \leq n_2$. Thus, $w(\varphi(k)) - w(k) \geq 1$ for $0 \leq k \leq n_1 - 1$ and

$$W(n_2+n_1-1) - W(n_2-1) = \sum_{k=n_2}^{n_2+n_1-1} w(k) = \sum_{k=0}^{n_1-1} w(\varphi(k))$$

$$\geq \sum_{k=0}^{n_1-1} (w(k) + 1) = W(n_1-1) + n_1$$

which is a contradiction to (10). Thus, Theorem 2 is proved.

Corollary 1: Let G be a primitive graph with n vertices. Then

$$|E(G)| \leq \left(\frac{n}{n-2}\right)W(n-2).$$

Proof: We have seen that $n \geq 3$. Suppose there exists a graph G with n vertices and $|E(G)| > \left(\frac{n}{n-2}\right)W(n-2)$. Then

$$|E(G)| > W(n-2) + \frac{2|E(G)|}{n}. \qquad (13)$$

Since G has n vertices and $|E(G)|$ edges then some vertex $v_0 \in V(G)$ must have degree $\leq \frac{2|E(G)|}{n}$. Form the graph G′ by removing v_0 from V(G) and all the edges of E(G) which are incident to v_0. Hence,

$$|V(G')| = n - 1$$

and

$$|E(G')| \geq |E(G)| - \frac{2|E(G)|}{n} > W(n-2) \qquad (14)$$

by (13). Since G is primitive by hypothesis then G′ must be completely decomposable. This is impossible however since (14) contradicts Theorem 2. This completes the proof.

V. SOME AUXILIARY RESULTS

We shall need several additional facts.

Lemma 3: For fixed integers $n_k \geq 1$, $1 \leq k \leq t$, let $(c_{i_1 \ldots i_t})_{1\leq i_k \leq n_k, 1\leq k\leq t}$ be a t-dimensional array of real numbers. Then the following two statements are equivalent:

(i) For all choices of $0 \leq x_1^{(k)} \leq x_2^{(k)} \leq \ldots \leq x_{n_k}^{(k)}$, $1 \leq k \leq t$,

$$\sum_{\substack{1\leq i_k\leq n_k \\ 1\leq k\leq t}} c_{i_1 \ldots i_t} x_{i_1}^{(1)} \ldots x_{i_t}^{(t)} \geq 0. \tag{15}$$

(ii) For all choices of $1 \leq j_k \leq n_k$, $1 \leq k \leq t$,

$$\sum_{\substack{j_k\leq i_k\leq n_k \\ 1\leq k\leq t}} c_{i_1 \ldots i_t} \geq 0. \tag{15′}$$

Proof: Suppose (ii) fails. Then for some choice of $1 \leq j_k \leq n_k$, $1 \leq k \leq t$, the sum in (15′) is < 0. Choose $x_1^{(k)} = \ldots = x_{j_k-1}^{(k)} = 0$, $x_{j_k}^{(k)} = \ldots = x_{n_k}^{(k)} = 1$, $1 \leq k \leq t$. A straightforward substitution into the sum of (15) shows that (i) also fails.

Suppose (ii) holds. The following identity is easily verified:

$$\sum_{\substack{1\leq i_k\leq n_k \\ 1\leq k\leq t}} c_{i_1 \ldots i_t} x_{i_1}^{(1)} \ldots x_{i_t}^{(t)}$$

$$= \sum_{\substack{1\leq j_k\leq n_k \\ 1\leq k\leq t}} \left(\sum_{\substack{j_k\leq i_k\leq n_k \\ 1\leq k\leq t}} c_{i_1 \ldots i_t} \right) \left(x_{j_1}^{(1)} - x_{j_1-1}^{(1)} \right) \ldots \left(x_{j_t}^{(t)} - x_{j_t-1}^{(t)} \right) \tag{16}$$

where $x_0^{(k)} = 0$, $1 \leq k \leq t$. From (16) it follows at once that (i) holds. This proves the lemma.

We state the simple special case $t = 1$ as

Corollary 2: Let $(c_i)_{1 \leq i \leq n}$, be a sequence of real numbers. Then $\sum_{1 \leq i \leq n} c_i x_i \geq 0$ for all choices of $0 \leq x_1 \leq x_2 \leq \ldots \leq x_n$ if and only if $\sum_{j \leq i \leq n} c_i \geq 0$ for all choices of $1 \leq j \leq n$.

Finally, we consider a set X and a pairing P of X with partition $X = X_0 \cup X_1$ and map $\theta : X_0 \to X_1$ (cf. Sec. III). Let f map the real line $\mathbb{R}$ into $\mathbb{R}$ such that f is symmetric, i.e., $f(x) = f(-x)$ for $x \in \mathbb{R}$. By the f-value of P, denoted by $v_f(P)$, we mean

$$v_f(P) = \sum_{x \in X_0} f(x - \theta(x))$$

(compare with (3)).

Lemma 4: Suppose f is symmetric, and for $x \geq 0$ convex and nondecreasing. Then for a fixed partition of $X = X_0 \cup X_1$, $v_f(P)$ is minimized by taking the map θ to be order-preserving.

Proof: Suppose there exist $x, x' \in X_0$ such that $x < x'$ and $\theta(x) > \theta(x')$. (There is no problem if $x = x'$ or $\theta(x) = \theta(x')$.) We show that if θ is replaced by $\theta' : X_0 \to X_1$ defined by

$$\theta'(x) = \theta(x'),$$

$$\theta'(x') = \theta(x),$$

$$\theta'(y) = \theta(y), \quad y \neq x, x',$$

then the resulting pairing P′ has $v_f(P') \leq v_f(P)$. It is sufficient to show

$$f(\theta(x') - x) + f(\theta(x) - x') \leq f(\theta(x) - x) + f(\theta(x') - x'). \tag{17}$$

By the symmetry of f we can assume $x \leq \theta(x')$. There are three cases.

(i) $x \leq x' \leq \theta(x') \leq \theta(x)$. Then (17) holds if

$$f((\theta(x')-x') + (x'-x)) + f((\theta(x)-\theta(x')) + (\theta(x')-x'))$$
$$\leq f((\theta(x)-\theta(x')) + (\theta(x')-x') + (x'-x)) + f(\theta(x')-x')$$

and this is a direct consequence of the convexity of f.

(ii) $x \leq \theta(x') \leq x' \leq \theta(x)$. Then (17) holds if

$$f(\theta(x')-x) + f(\theta(x)-x')$$
$$\leq f(x'-\theta(x')) + f((\theta(x)-x') + (x'-\theta(x')) + (\theta(x')-x)).$$

But convexity implies

$$f(\theta(x')-x) + f(\theta(x)-x') \leq f(0) + f((\theta(x')-x) + (\theta(x)-x'))$$

and by monotonicity we can add $x' - \theta(x')$ to each of the arguments to establish (17) in this case.

(iii) $x \leq \theta(x') \leq \theta(x) \leq x'$. Then (17) holds if

$$f(\theta(x')-x) + f(x'-\theta(x)) \leq f((\theta(x)-\theta(x')) + (\theta(x')-x))$$
$$+ f((\theta(x)-\theta(x') + (x'-\theta(x)))$$

and this is a consequence of the monotonicity of f.

This proves the lemma.

VI. OPTIMAL COMPLETE PAIRINGS

We recall the notation of Sec. III. Let $A = \{a_0 \leq \cdots \leq a_{2^n-1}\}$. The main purpose of this section is to determine $\mu(A) \equiv$ minimum of $V(P^*)$ over all complete pairings P^* of A.

Theorem 3:

$$\mu(A) = \nu(A) = \sum_{k=0}^{2^n-1} (2w(k) - n)a_k.$$

Proof: Let P^* be a complete pairing of A. Then

$$V(P^*) = \sum_{k=0}^{2^n-1} I_k a_k \tag{18}$$

where the I_k are integers which are $\equiv n \pmod 2$. This follows by observing that $V(P^*)$ is just a sum of terms of the form $a_i - a_j$, $i > j$ (since we sum the absolute value of the differences). Thus, each a_k occurs with a coefficient of ± 1 and altogether there are n coefficients to add up for each a_k. This sum we denote by I_k; this is (18). Specifically I_k is the number of terms $a_k - a_i$ in which a_k was paired with a smaller a_i, minus the number of terms $a_j - a_k$ in which a_k was paired with a larger a_j.

Now consider the following complete pairing Q* of A. All partitions of various subsets $B = \{b_0 \leq \cdots \leq b_{2^r-1}\} \subseteq A$ are of the form $B = B_0 \cup B_1 = \{b_0 \leq \cdots \leq b_{2^{r-1}-1}\} \cup \{b_{2^{r-1}} \leq \cdots \leq b_{2^r-1}\}$; all maps $\theta : B_0 \to B_1$ are order-preserving. By Lemma 4 once we have made the partition into B_0 and B_1, we may as well take θ to be order-preserving in order to minimize V.

This complete pairing will be called the <u>canonical</u> complete pairing. It is easy to verify

$$V(Q^*) = \nu(A). \tag{19}$$

In Q*, each a_k is paired with all a_i for which $|w(k) - w(i)| = 1$. Hence, we can change any one of the $w(k)$ 1's to a 0 in the binary expansion of k to obtain an admissible i or we can change any one of the $n - w(k)$ 0's to a 1. This implies the coefficient of a_k is $w(k) - (n - w(k)) = 2w(k) - n$ and

$$V(Q^*) = \sum_{k=0}^{2^n-1} (2w(k) - n)a_k. \tag{20}$$

Suppose now that P* is a complete pairing of A such that $V(P^*) < V(Q^*)$. Then

$$\begin{aligned} &\sum_{k=0}^{2^n-1} I_k a_k < \sum_{k=0}^{2^n-1} (2w(k) - n)a_k, \\ &\sum_{k=0}^{2^n-1} (I_k - 2w(k) + n)a_k < 0. \end{aligned} \tag{21}$$

By Corollary 2, (21) implies that for some r, $0 \leq r \leq 2^n - 1$,

$$\sum_{k=r}^{2^n-1} (I_k - 2w(k) + n) < 0,$$

$$\sum_{k=r}^{2^n-1} (n - w(k)) < \sum_{k=r}^{2^n-1} \left(\frac{n-I_k}{2}\right),$$

$$\sum_{k=r}^{2^n-1} \left(\frac{n-I_k}{2}\right) > \sum_{k=r}^{2^n-1} (n - w(k)) = \sum_{k=r}^{2^n-1} w(2^n-1-k)$$

$$= \sum_{k=0}^{2^n-r-1} w(k) = W(2^n-r-1). \qquad (22)$$

Notice that the term $\frac{n-I_k}{2}$ is just the number of coefficients of a_k in V which have the coefficient -1, i.e., the number of a_j paired with a_k for which $j > k$.

Let us associate a graph G with the complete pairing P* in the following way. V(G) will be the set $\{0,1,\ldots,2^n-1\}$; $\{i,j\} \in E(G)$ if and only if a_i was paired with a_j for some pairing in P*. Thus each vertex of G has degree n. Note that the graph associated with the canonical complete pairing Q* is just the n-cube C^n. We point out the following important

Fact 2: G is completely decomposable. This follows immediately from the construction of a complete pairing.

Consider the subgraph G′ of G defined by $V(G') = \{r,\ldots,2^n-1\}$ and $\{i,j\} \in E(G')$ if and only if $i,j \in V(G')$ and $\{i,j\} \in E(G)$. For each k, $r \leq k \leq 2^n - 1$, there is an edge in E(G′) for each $j > k$ such that a_j and a_k were paired in P*. Hence

$$|E(G')| = \sum_{k=r}^{2^n-1} \left(\frac{n-I_k}{2}\right). \qquad (23)$$

Therefore, by (23) and (22) and the definition of G′,

$$|E(G')| > W(2^n-r-1), \quad |V(G')| = 2^n - r. \qquad (24)$$

However, this is a contradiction to Theorem 2 since G is completely decomposable and therefore the subgraph G′ must also be completely decomposable. This shows

$$V(P^*) \geq V(Q^*) \tag{25}$$

which, together with (19) and (20), proves the theorem.

VII. CONCLUDING REMARKS

A number of open questions remain, several of which we list here.

1. Suppose the definition of v(P) in (3) is replaced by

$$v(P) = \sum_{x \in X_0} (x - \theta(x))^2. \tag{3'}$$

Is it true that for $A = \{0,1,\ldots,2^n-1\}$ the canonical pairing Q* still has the minimum V-value? It has been recently shown in [1] that in the case of cube-numbering, the canonical assignment is still optimal; it is reasonable to conjecture that Q* is the optimal complete pairing here. One may also use a more general f-value of P (cf. Lemma 4) and more general sets A and ask when Q* is optimal. For the related problem in which $v(P) = \min_{x \in X_0} |x - \theta(x)|$, this has been solved for $A = \{0,1,\ldots,2^n-1\}$ by Harper [3].

2. Suppose G is a primitive graph. Can $|E(G)|/|V(G)|$ be arbitrarily large? The bound in Corollary 1 only shows for $\varepsilon > 0$,

$$|E(G)|/|V(G)| < (\tfrac{1}{2} + \varepsilon) \frac{\log n}{\log 2}$$

for n sufficiently large. We can also ask what reasonable lower bounds may be placed on $|E(G)|/|V(G)|$. For example, it is not difficult to show that except for $G = K_3$, a triangle, any primitive graph G must satisfy $|E(G)|/|V(G)| \geq 6/5$.

3. Must all the edges of a primitive graph be regular? Must a primitive graph have a vertex of degree 2? No counterexamples are known at present.

4. Can a primitive graph G have an even number of vertices? While the answer to this question is almost certainly yes, it is known that there are no primitive graphs with 2, 4 or 6 vertices.

5. Can a primitive graph have smallest cycle length $\geq k$? At present, the answer is known to be in the

affirmative only for $k \leq 4$.

6. More generally, is it possible to classify the primitive graphs? Because of the rather strong conditions a primitive graph must satisfy, this goal may not be too unreasonable.

REFERENCES

1. Crimmins, T. R., Horwitz, H. M., Palermo, C. J. and Palermo, R. V., Minimization of Mean-Square Error for Data Transmitted by Group Codes, IEEE Trans. Inf. Theory, Vol. IT-15, No. 1 (1969) 72-78.

2. Harper, L. H., Optimal assignments of numbers to vertices, J. SIAM 12 No. 1 (1964) 131-135.

3. Harper, L. H., Optimal numberings and isoperimetric problems on graphs, J. Comb. Theory, 1 No. 3 (1966) 385-393.

4. Kautz, W. H., Optimized data encoding for digital computers, Convention Record I.R.E., (1954) 47-57.

5. Lindsey, J. H., Assignment of numbers to vertices, Amer. Math. Monthly 71 (1964) 508-516.

6. Ore, O., Theory of graphs, Amer. Math. Soc. Colloq. Pub. 38 (1962) Providence.

7. Steiglitz, K. and Bernstein, A. J., J. SIAM 13 No. 2 (1965) 441-443.