Group Codes for the Gaussian Channel

By DAVID SLEPIAN

(Manuscript received April 27, 1967)

A class of codes for use on the Gaussian channel, called group codes, is defined and investigated. Roughly speaking, all words in a group code are on an equal footing: each has the same error probability and the same disposition of neighbors. A decomposition theorem shows every group code to be equivalent to a direct sum of certain basic group codes generated by real-irreducible representations of a finite group associated with the code. Some theorems on distances between words in group codes are demonstrated. The difficult problem of finding group codes with large nearest neighbor distance is discussed in detail.

I. INTRODUCTION

In a communication model first introduced by Kotel'nikov[1] in 1947, and independently by Shannon[2] in 1948, and since studied by many authors,[3–22] messages for transmission are represented by vectors in a Euclidean space, $\mathcal{S}_n$, of n dimensions called signal space. In this model, known as the Gaussian channel, when $\mathbf{X}$ is transmitted, the received signal is represented by a vector $\mathbf{Z} = \mathbf{X} + \mathbf{Y}$ which consists of the sum of the sent vector and a noise vector $\mathbf{Y}$ whose components are independent Gaussian variates with mean zero and variance σ^2. Some physical circumstances that lead to this model, as well as further details, can be found in Refs. 3, 10, and 13.

An equal-energy block code of size M for use on this Gaussian channel is a collection of M distinct vectors $\mathbf{X}_1, \mathbf{X}_2, \cdots, \mathbf{X}_M$ in signal space all of the same length. We shall always suppose $M \geqq n$ and that the vectors span $\mathcal{S}_n$. The length of the vectors serves to define an important parameter S called the average power of the code through the equation

$$nS = |\mathbf{X}_i|^2. \tag{1}$$

The vectors of the code are called code words or code points. Their termini lie on the sphere of radius $\sqrt{nS}$ centered at the origin of $\mathcal{S}_n$.

Associated with each code point $\mathbf{X}_i$ of an equal-energy block code

is a region $\mathcal{R}_i$ of signal space called a maximum likelihood region and defined by

$$\mathcal{R}_i = \left\{ \mathbf{X} \;\middle|\; | \mathbf{X} - \mathbf{X}_i | \leqq | \mathbf{X} - \mathbf{X}_j |, j \neq i \right\}. \tag{2}$$

That is, $\mathcal{R}_i$ is the set of all points in $\mathcal{S}_n$ at least as close to $\mathbf{X}_i$ as to any other code word. These regions are convex flat-sided cones with apex at the origin. The interiors of $\mathcal{R}_i$ and $\mathcal{R}_j$ are disjoint for $i \neq j$: the union of the $\mathcal{R}_i$ is all of $\mathcal{S}_n$.

The capabilities of equal energy block codes for communicating over the Gaussian channel are well known. If the words of a code are presented equally likely and independently for transmission over the channel, the communication rate is

$$R = \frac{\alpha}{n} \log M \tag{3}$$

natural units per second where α (measured in numbers per second) is the rate at which vector components are transmitted. The receiver which minimizes the average error probability[5,13] operates by asserting that code word $\mathbf{X}_i$ was transmitted when the received vector $\mathbf{Z}$ lies in $\mathcal{R}_i$, $i = 1, 2, \cdots, M$. (The received vector lies in the boundary of some $\mathcal{R}_i$ with probability 0.) When $\mathbf{X}_i$ was transmitted the error probability of this best receiver is

$$P_{ei} = \frac{1}{(2\pi\sigma^2)^{n/2}} \int_{\mathcal{R}_i'} \cdots \int \exp\left(-\frac{1}{2\sigma^2} |\mathbf{Y} - \mathbf{X}_i|^2 \right) dy_1 \cdots dy_n \tag{4}$$

where $\mathcal{R}_i'$ is the complement of $\mathcal{R}_i$. The average error probability is

$$P_e = \frac{1}{M} \sum_{i=1}^{M} P_{ei}. \tag{5}$$

Upper and lower bounds are known[4,6,7,11,17] for $P_{e\,\min}(M, n, S)$, the smallest attainable value of P_e for an equal-energy block code with the indicated parameters. In the limit as $n \to \infty$, these bounds lead to the famous capacity formula $C = \alpha/2 \log (1 + S/\sigma^2)$ whose interpretation we suppose known. For fixed finite values of M and n, however, little is known in the general case about codes for which P_e attains its minimal value (optimal codes). The cases $M = n + 1, n + 2, \cdots, 2n$ have been studied in some detail.[8,9,14] For $n = 2$, Weber[14] showed that the regular M-gon is globally optimal: for $M = n + 1$, $n = 2, 3, \cdots$, it has been shown[20] that the regular simplex is optimal. No other optimal codes with $n > 3$ are known.

Recently Wyner[12] has investigated the capabilities of equal-energy block codes when a suboptimal receiver, known as a bounded distance decoder, is used. Here the regions $\mathcal{R}_i$ of the maximum likelihood receiver are replaced by spheres of radius $d/2$ centered on the termini of the code vectors $\mathbf{X}_i$, where d is the minimum distance between any two words of the code. If the received vector is not in one of these spheres, a decoding error is assumed. Wyner established upper and lower bounds on the smallest error probability attainable with an equal-energy block code using bounded distance decoding. In the limit as $n \to \infty$ he obtained coding theorems and a capacity analogous to the usual ones. For finite M and n, the error probability using bounded distance decoding is a monotone decreasing function of the minimum distance d between code words of an equal-energy block code. In the general case little is known about equal-energy block codes with largest nearest neighbor distance.

For equal energy block codes of M vectors spanning $\mathcal{S}_n$ two optimization problems thus present themselves: to find a code for which P_e, as given by (4) and (5), is a minimum; and to find a code with largest nearest neighbor distance between its code words. We have made little progress in solving these problems.

In this paper we investigate instead a class of equal-energy block codes called group codes. It is conjectured that this class includes solutions to the problems just mentioned for many values of M and n. Quite apart from these questions of optimality, however, group codes possess an important symmetry property that makes their study of interest in its own right. Roughly speaking, all code words in a group code are on an equal footing. This notion is made precise in the next section.

Most codes that have been investigated for the Gaussian channel are group codes: it is likely that any code used in practice will be of this type. Group codes for the Gaussian channel are a natural extension of the group codes introduced for the binary channel in Ref. 21, and these latter codes are obtained as a special case of the codes described here.

In what follows, we define equivalence for group codes, then investigate the possible classes of group codes. Here the theory of group representations plays a key role.[25] The appendix gives a summary of results needed from this field. The problem of constructing group codes is considered and an optimization problem of some difficulty is encountered. A number of interesting properties of group codes are disclosed.

Many of the results reported here are contained in the author's Bell Telephone Laboratories report of May 7, 1951, a document that received a limited circulation outside the Laboratories. A number of these results were recently rediscovered independently by J. G. Dunn and appear in his thesis[22] along with extensions in directions different from those reported here. The discovery of an easy decoding algorithm for certain group codes[15] has led to a revival of the author's interest in this subject, and so the present paper, while in part very old, is a report on research now in progress. It examines the general structure of group codes. In a later paper we hope to give a detailed treatment of some group codes associated with the symmetric group.

II. GROUP CODES

In studying the geometric properties of equal-energy block codes, it is convenient to deal only with code vectors of unit length. That is, we set S in equation 1 equal to $1/n$, and deal with normalized codes. To compute error probabilities associated with the use of the code, one must scale up the vectors by a factor $\sqrt{nS}$.

Let $\mathbf{X}_1, \mathbf{X}_2, \cdots, \mathbf{X}_M$ be the (unit) vectors of an equal-energy block code. It is clear from the definition of the regions $\mathfrak{R}_i$ and from (4) and (5) that P_e is invariant under a rotation of the code as a whole. That is, if O is an arbitrary $n \times n$ orthogonal matrix and

$$\mathbf{X}_i' = O\mathbf{X}_i, \qquad i = 1, 2, \cdots, M, \tag{6}$$

the error probability P_e' for the code $\mathbf{X}_1', \cdots, \mathbf{X}_M'$ is the same as that for the code $\mathbf{X}_1, \mathbf{X}_2, \cdots \mathbf{X}_M$. The set of interword distances for the two codes is the same, and in particular both codes have the same minimum nearest neighbor distance d. Two codes whose vectors (with possible renumbering) can be related as in equation (6) are called *equivalent*. Equivalent codes have the same communication capabilities.

We now examine in what sense the words of an equal-energy block code in S_n might be "alike". Given the M unit vectors $\mathbf{X}_i$ that define the code, the real orthogonal n by n matrix O is said to leave the code invariant if the $\mathbf{Y}_i$ are a permutation of the $\mathbf{X}_i$ where $\mathbf{Y}_i = O\mathbf{X}_i$, $i = 1, 2, \cdots, M$. The collection $\theta = \{O_1, O_2, \cdots, O_g\}$ of all real orthogonal n by n matrices that leave the code invariant clearly forms a finite* group under ordinary matrix multiplication. Now transformation by

* By hypothesis, the $\mathbf{X}_i$ span S_n. An $n \times n$ orthogonal matrix is completely determined by its effect on a set of n vectors that span its carrier space. Since the words of the code are permuted along themselves by each element of θ, $g \leqq M!$.

an orthogonal matrix preserves distances between points, so that a possible definition of "alikeness" for the points of the code is to require that in the group θ there be elements O_1 , O_2 , $\cdots$, O_M that transform any particular word, say $\mathbf{X}_i$, into each of the M vectors of the code. A collection of M unit vectors spanning S_n that satisfies this condition will be called a *group code* and denoted by the symbol $\{M, n\}$. In a group code, if O_i sends $\mathbf{X}_1$ into $\mathbf{X}_i$ and O_j sends $\mathbf{X}_1$ into $\mathbf{X}_j$, then $O_jO_i^{-1}$ sends $\mathbf{X}_i$ into $\mathbf{X}_j$. We have then

Proposition 1: For a group code, the set of distances from $\mathbf{X}_i$ *to all other points of the code is the same as the set of distances from* $\mathbf{X}_j$ *to all other points of the code,* $i, j = 1, 2, \cdots, M$.

Each point has the same number of nearest neighbors, the same number of next nearest neighbors, and so on.

The maximum likelihood regions $\mathfrak{R}_i$ for a code are defined by equation (2) in terms of distances from code points. Since orthogonal matrices leave distances invariant, it follows that for a group code a matrix $O \in \theta$ that sends $\mathbf{X}_i$ into $\mathbf{X}_j$ also sends $\mathfrak{R}_i$ into $\mathfrak{R}_j$. From this fact and the form of (4) we have

Proposition 2: For a group code $\{M, n\}$ *the maximum likelihood regions* $\mathfrak{R}_1$, $\mathfrak{R}_2$, $\cdots$, $\mathfrak{R}_M$ *are all congruent and all words have the same error probability, that is,* $P_{e1} = P_{e2} = \cdots = P_{eM} = P_e$.

III. GENERATION AND CLASSIFICATION OF GROUP CODES

To each matrix O of the group θ of orthogonal matrices that leaves a group code $\{M, n\}$ invariant, there corresponds a permutation on M letters, namely the permutation effected by O on the M vectors of the code. That these permutations form a transitive permutation group follows from the definition of a group code. No two different elements of θ can effect the same permutation of the words of $\{M, n\}$ since the effect of an $n \times n$ matrix on a set of vectors spanning S_n completely determines the matrix. We have then

Proposition 3: The group θ *of all orthogonal* $n \times n$ *matrices leaving a group code* $\{M, n\}$ *invariant forms a faithful representation of (is simply isomorphic to) a transitive permutation group on* M *letters.*

Group codes $\{M, n\}$ do not exist for every M and n. For example, it is not hard to prove that it is impossible to arrange 5 points on the sphere in 3 dimensions to form a group code. Necessary and sufficient conditions on M and n for the existence of an $\{M, n\}$ are not known.

Group codes do exist in great abundance, however, and we shall give examples later. Indeed, from any set of $n \times n$ orthogonal matrices $O_1, O_2, \cdots, O_M$ that form a finite group $\mathcal{G}$ under matrix multiplication we can form a group code by choosing a unit n-vector $\mathbf{X}$ and forming the set of vectors

$$\mathbf{X}_i = O_i\mathbf{X}, \qquad i = 1, 2, \cdots, M. \tag{7}$$

Elements of $\mathcal{G}$ leave this configuration of vectors invariant by the group property. Since $\mathcal{G}$ must contain the $n \times n$ unit matrix, $\mathbf{X}$ is among the collection of vectors and it is sent into each of the other vectors. A group code therefore results. This code may not have M distinct vectors, however, and it may not span $\mathcal{S}_n$. The code depends on the initial vector $\mathbf{X}$.

If the code has fewer than M vectors, then for some $i \neq j$, $\mathbf{X}_i = \mathbf{X}_j$ or $O_i\mathbf{X} = O_j\mathbf{X}$, or $O_j^{-1}O_i\mathbf{X} = O_k\mathbf{X} = \mathbf{X}$ for some $O_k \in \mathcal{G}$. That is, $\mathbf{X}$ must be an eigenvector with eigenvalue unity for at least one $O \in \mathcal{G}$ different from the unit matrix. The set of all such $O \in \mathcal{G}$ forms the subgroup $\mathcal{H}$ of order h of $\mathcal{G}$ that sends $\mathbf{X}$ into itself. It is easy to show that by (7) $\mathcal{G}$ generates $\nu = M/h$ distinct vectors. Since the matrices of $\mathcal{G}$ have only a finite number of eigenvectors, however, it is always possible to choose an $\mathbf{X}$ so that the M vectors (7) are distinct.

It may not be possible, however, to choose $\mathbf{X}$ so that the vectors span $\mathcal{S}_n$. To discuss this matter further we must recall the notion of real-reducibility. A finite group of (real) orthogonal matrices $\mathcal{G} = O_1, O_2, \cdots, O_M$ is said to be real-reducible if there exists an $n \times n$ real orthogonal matrix O such that for $i = 1, 2, \cdots, M$

$$OO_iO^{-1} = \left(\begin{array}{c|c} A_i & D \\ \hline C & B_i \end{array}\right) \tag{8}$$

where A_i is an l by l matrix, B_i is an $n - l$ by $n - l$ matrix, $0 < l < n$ and C and D are matrices all of whose elements are zero. It is assumed that l does not depend on i. A group of real orthogonal matrices that is not real-reducible is said to be real-irreducible. In words, a real-reducible collection of matrices can be simultaneously transformed to block diagonal form by a real orthogonal matrix: a real-irreducible collection cannot be so reduced.* The reduced matrix in block form in equation (8) is said to be the direct sum of the two square matrices A_i and B_i.

* In the theory of group representations (see the appendix) reducibility is usually defined over the field of complex numbers. The definition is as above with O replaced by a *unitary* matrix. We shall speak simply of "reducibility" in this case as opposed to "real-reducibility".

It is easy to show that if the matrices O_i of equation (7) are real-irreducible, then the code they generate spans $\mathcal{S}_n$ for all choices of $\mathbf{X}$: if they are real-reducible, for some choices of $\mathbf{X}$ the code will not span $\mathcal{S}_n$.

These comments lead to

Proposition 4: Every real-irreducible group $\mathcal{G} = O_1, O_2, \cdots, O_M$ *of real orthogonal* $n \times n$ *matrices serves by means of equation* (7) *to generate a group code* $\{M', n\}$ *for each unit vector* $\mathbf{X}$ *in* $\mathcal{S}_n$. *Here* $M' \leqq M$. *If* $M' < M$, *it is a divisor of* M.

Propositions 3 and 4, together with the theory of group representations† suggest a means of classifying and generating all group codes. From Proposition 3 we can associate with a given group code $\{M, n\}$ a unique abstract group and a faithful representation θ of this group by orthogonal matrices. The code can be thought of as generated from one of its vectors, $\mathbf{X}$, say, by the operation of the matrices of this representation in the manner of equation (7). Now the representation θ will in general be real-reducible. There will exist then a real orthogonal matrix O that will exhibit θ in block form (8) as the direct sum of a number of real-irreducible representations. Denote this new reduced representation by θ'. It is easily seen that the matrices of θ' operating on the vector $\mathbf{Y} = O\mathbf{X}$ generate a group code $\mathbf{Y}_1, \mathbf{Y}_2, \cdots, \mathbf{Y}_M$ equivalent to the originally given $\{M, n\}$. We can further regard $\mathbf{Y}$ as the sum of its projections $\mathbf{Y}^1, \mathbf{Y}^2, \cdots$ on the various invariant subspaces of θ' indicated by its block structure.

By the procedure just outlined, for each equivalence class of group codes we arrive at a particular set of real-irreducible representations, say $\theta_1, \theta_2, \cdots, \theta_j$ of an abstract group, each with a corresponding associated vector $\mathbf{Y}^1, \mathbf{Y}^2, \cdots, \mathbf{Y}^j$. We regard $\mathbf{Y}^i$ as lying in the carrier space of θ_i, so that if θ_i is of dimension l_i, then $\mathbf{Y}^i$ is a vector of l_i components, $i = 1, 2, \cdots, j$. Let the length of $\mathbf{Y}^i$ be λ_i. We have $\sum \lambda_i^2 = 1$. The θ_i are determined by $\{M, n\}$ only up to equivalence in the sense of representation theory, owing to the possibility of reduction of θ by different matrices O. The vectors $\mathbf{Y}^i$ inherit some additional freedom owing to the M possible choices of $\mathbf{X}$ in the preceding paragraph.

We can think of the $\{M, n\}$ as decomposed by the above process into an equivalent direct sum of j group codes, the ith code being generated by the matrices of θ_i operating on the initial unit vector $\mathbf{Z}^i = \mathbf{Y}^i/\lambda_i$, $i = 1, 2, \cdots, j$. The constituent group codes are weighted by the numbers $\lambda_1, \lambda_2, \cdots, \lambda_j$ in forming the direct sum code $\{M, n\}$. Notice

† Knowledge of the material in the appendix is necessary for understanding much of the remainder of this paper.

that some of the constituent codes may have fewer than M distinct words.

Conversely, within equivalence we can construct any group code as the weighted direct sum of codes generated by real-irreducible groups of matrices after the manner of Proposition 4. In synthesizing codes in this manner, we may, of course, arrive at equivalent codes by several different constructions. The group $\mathcal{G}$ of Proposition 4 may be only a subgroup of the group of all orthogonal transformations that leave the code generated by $\mathcal{G}$ invariant. Different initial vectors operated on by the same group of matrices may give rise to equivalent codes.

Every group possesses the trivial real-irreducible one-dimensional *identity representation* in which each group element is represented by the one-dimensional unit matrix. The inclusion of this identity representation in the constituent codes making up a direct sum code represents a waste of one dimension, since the code is then equivalent to one in which each code vector has the same first component. This first component then carries no information. By omitting the first component of each vector (and rescaling the length of the resultant vectors), a new code of dimension $n - 1$ is obtained with error probability no greater than the original $\{M, n\}$. In general in what follows we will not be concerned with codes that contain this identity representation.

We turn our attention now to the basic problem of constructing good group codes as the weighted sum of properly chosen group codes generated by real-irreducible groups of orthogonal matrices.

IV. THE INITIAL VECTOR PROBLEM AND THE FUNDAMENTAL REGION

As in Proposition 4, let a code be constructed from a given group $\mathcal{G} = O_1, O_2, \cdots, O_M$ of orthogonal $n \times n$ matrices by means of equation (7). We think of these matrices as a faithful representation of an abstract group isomorphic to the matrix group. The code obtained in this manner depends upon the initial $\mathbf{X}$ on which the matrices operate. The regions $\mathcal{R}_i$ of equation (2) and hence also $P_e = P_{e_i}$ by (4) also depend on this choice. We suppose now that $\mathbf{X}$ is not an eigenvector of any of the O_i so that the code has M distinct words. It would be desirable to be able to choose an $\mathbf{X}$ of this sort to either minimize P_e or to maximize d, the nearest neighbor distance. We have not seen how to solve either of these problems in general. A few words about them are in order.

Consider first the problem of choosing $\mathbf{X}$ to maximize d. The squared

distance between $\mathbf{X}$ and $\mathbf{X}_i$ is

$$d^2(\mathbf{X}, \mathbf{X}_i) = |\mathbf{X} - \mathbf{X}_i|^2 = 2 - 2\mathbf{X}\cdot O_i\mathbf{X}$$

a monotone decreasing function of the quadratic form $\mathbf{X}\cdot O_i\mathbf{X}$ in the components of $\mathbf{X}$. This form is the cosine of the angle between $\mathbf{X}$ and $\mathbf{X}_i$. Solution of the maximum nearest neighbor distance is equivalent to finding

$$\alpha = \min_{\mathbf{X}} \max_{i} \mathbf{X}\cdot O_i\mathbf{X} \tag{9}$$

where the maximization over the matrices of $\mathcal{G}$ must omit the identity matrix. The quantity α is an invariant of the representation (is the same for every equivalent representation) and should ultimately be expressible in terms of properties of the group. The vector $\mathbf{X}$ which minimizes (9) is not unique: any word in the code generated by $\mathbf{X}$ would serve as well.

Given $\mathcal{G}$, we define two points $\mathbf{X}$ and $\mathbf{Y}$ on the unit sphere to be equivalent if one can be obtained from the other by an operation of $\mathcal{G}$. The surface of the sphere is thus divided into equivalence sets. A connected region on the sphere such that no two points in its interior are equivalent and such that every point on the sphere is equivalent to some point in the region will be called a fundamental region of $\mathcal{G}$. The maximum likelihood regions, $\mathcal{R}_i$, associated with any $\{M, n\}$ generated by $\mathcal{G}$ intersect the unit sphere in fundamental regions. These intersections are very special fundamental regions: they are convex and bounded by hyperplanes.

In attempting to minimize P_e or maximize d it clearly suffices to consider initial vectors $\mathbf{X}$ restricted to some fundamental region. It is natural then to ask what fundamental regions are possible for a given $\mathcal{G}$.

The situation is complicated. For some groups, the fundamental region is completely determined (up to equivalence under the group operations, of course): for other groups only certain features of its boundaries are determined, or no points at all may be determined.

For example, in the plane consider the group $\mathcal{G}_1$ generated by the three matrices corresponding to reflections in three lines through the origin that make angles of 60° with each other. This group is of order 6 and is a subgroup of the symmetry group of a regular hexagon having the given lines as diagonals. The fundamental region of this group is completely determined. It is a 60° arc of the unit circle with end points on two of the given lines. Any group code $\{6, 2\}$ generated by $\mathcal{G}_1$ has

this fundamental region for the intersection of one of its maximum likelihood regions $\mathcal{R}_i$ with the circle. Choice of **X** serves only to position the initial vector within the maximum likelihood region. (When **X** is chosen to lie on one of the reflection lines, a $\{3, 2\}$ results and the maximum likelihood region changes discontinuously to the union of two adjacent regions of the sort just discussed.)

On the other hand, consider the group $\mathcal{G}_2$ of rotational symmetries of the regular hexagon. $\mathcal{G}_2$, of order 6, consists of a 2×2 matrix representing a rotation of 60° in the plane along with the distinct powers of this matrix. *Any* 60° arc of the unit circle is a fundamental region for this group. Codes $\{6, 2\}$ generated by $\mathcal{G}_2$ are equivalent for all choices of the initial vector **X**.

An example illustrating a partly determined fundamental region is obtained by considering the pure rotational symmetries of a cube in three dimensions. We imagine the cube centered at the origin and inscribed in a unit sphere. We speak in terms of the operations on the cube rather than in terms of the 3×3 matrices which describe these operations. $\mathcal{G}_3$, a group of order 24, consists of rotations of the cube by 120° around the body diagonals, of rotations by 90° about axes through the origin and centers of faces and of rotations of 180° about axes through the midpoints of edges and the origin. One axis of each kind is shown on Fig. 1. In discussing the fundamental region of $\mathcal{G}_3$ and codes generated by $\mathcal{G}_3$, it is convenient to speak of points on the cube, rather than on the circumscribed unit sphere. It is to be understood

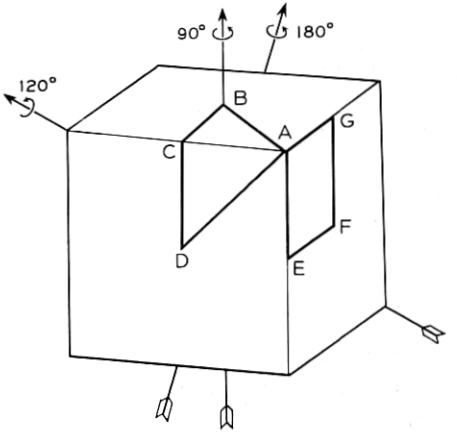

Fig. 1 — Example of partly-determined fundamental region.

then that when a point on the cube is mentioned it is really the corresponding point on the sphere obtained by projecting along a radius that is under discussion.

The vertices of the cube, the centers of faces and the midpoints of edges must all lie in boundaries of fundamental regions, for these points are on axes of rotation of $\mathcal{G}_3$. For example, a point distance ϵ from a vertex of the cube has two nearby equivalent points forming an equilateral triangle with the vertex at the center of the triangle. These three points cannot all lie in the interior of one fundamental region. The cube vertex therefore cannot be an interior point of a fundamental region. In fact at least 3 fundamental regions must meet at each vertex, at least 4 at each face center and at least two at each edge midpoint. Cube vertices and face centers must therefore be vertices of fundamental regions. Now all vertices of the cube are equivalent under $\mathcal{G}_3$ as are all face centers and all edge midpoints; no two of these three types of points are equivalent. A fundamental region of $\mathcal{G}_3$ must therefore contain at least one cube vertex and one face center among its vertices and at least one cube edge midpoint along its boundary.

Two distinct types of fundamental regions for $\mathcal{G}_3$ bounded by hyperplanes (great circles on the sphere) are shown in Fig. 1. Region *AEFG* is bounded by four hyperplanes. Edge midpoints are vertices of this type or region. Four fundamental regions surround each face center and each edge midpoint: three surround each cube vertex. Region *ABCD* is bounded by only three hyperplanes. Edge midpoints are no longer vertices of the fundamental region. Eight regions meet at each face center. The fundamental region *ABCD* corresponds to the maximum likelihood region of a group code having an initial vector (and hence all vectors) pass through a cube edge: region *AEFG* results when the initial vector passes through a face diagonal. All other positions of the initial vector give maximum likelihood regions that are fundamental regions bounded by four hyperplanes but not congruent to *AEFG*.

$\mathcal{G}_3$ is the irreducible representation of the symmetric group on four letters derived from the Young tableau[25] associated with the partition (2, 1, 1). The irreducible representation belonging to the partition (3, 1) is also three dimensional. It is equivalent to the group of symmetries of the regular tetrahedron and can be generated by reflections in planes through the centroid of the tetrahedron and its edges. The fundamental region here is completely determined. It is bounded by three of these generating reflection planes. Maximization of nearest neighbor distance for a {24, 3} generated by this group can be easily accomplished by

choosing the initial vector equidistant from the three bounding planes of the fundamental region.

More generally, Coxeter[23] has shown that if a real-irreducible finite group of $n \times n$ orthogonal matrices is generated by reflections, the fundamental region is completely determined, and in fact the region is bounded by n hyperplanes. Indeed, Coxeter has enumerated all possible groups of this sort. In dimensions n greater than 8, there are only three such groups, called by him A_n, B_n, and C_n of order $(n+1)!$, $2^{n-1}n!$, and $2^n n!$, respectively. These groups generate permutation modulation codes[15]—A_n generates Variant I codes, B_n generates Variant II codes with $\mu_1 = 0$, and C_n generates Variant II codes with $\mu_1 \neq 0$. The various permutation modulation codes are obtained by choosing the initial vector to lie in boundaries of various dimensionality of the fundamental regions of these groups.

Returning to the general case (when $\mathcal{G}$ is not generated by reflections), the real eigenvectors of the O_i with eigenvalue unity serve to determine landmarks of the fundamental region. Such an eigenvector must lie in the boundary of the region. If O_i has l such eigenvectors, their span is an l-dimensional boundary of the fundamental region. The situation has been studied by Robinson[24] in some detail, but no simple method of classifying the possible regions is available.

V. THE DIRECT SUM

Since any group code is equivalent to the weighted direct sum of codes generated by real-irreducible representations of a group, it is natural to investigate the relationship between interword distances in the sum code and the corresponding distances in the summand codes.

Let $\mathcal{G} = A_1, A_2, \cdots, A_g$ be a finite group of order g with A_1 the identity. Let $D^1(A)$ and $D^2(A)$ be two real-irreducible representations of $\mathcal{G}$ by real orthogonal matrices of dimensions l_1 and l_2 respectively. Let $\mathbf{X}_i = D^1(A_i)\mathbf{X}$, and $\mathbf{Y}_i = D^2(A_i)\mathbf{Y}$, $i = 1, 2, \cdots, g$ be group codes generated by D^1 and D^2. (Neither code need have g distinct vectors.) The direct sum code with weights λ_1 and λ_2 has vectors

$$\mathbf{Z}_i = \lambda_1 \mathbf{X}_i \oplus \lambda_2 \mathbf{Y}_i \qquad i = 1, 2, \cdots, g \tag{10}$$

$$\lambda_1^2 + \lambda_2^2 = 1, \qquad 0 < \lambda_1, \lambda_2 < 1$$

of $l = l_1 + l_2$ components. We seek to choose the weights so that the nearest neighbor distance, d, for the sum code $\mathbf{Z}$ is a maximum.

Let $\alpha_i = d^2(\mathbf{X}_i, \mathbf{X}_1)$ and $\beta_i = d^2(\mathbf{Y}_i, \mathbf{Y}_1)$ be the squared distance from the code word generated by A_i to the initial vector in the two codes,

$i = 1, 2, \cdots, g$, respectively. For the sum code we have

$$d^2(\mathbf{Z}_i, \mathbf{Z}_1) = \lambda_1^2\alpha_i + \lambda_2^2\beta_i$$

since the subspaces containing the **X** code and the **Y** code are orthogonal. The desired maximum nearest neighbor distance is thus

$$d^2 = \max_{0 \leqq \lambda \leqq 1} \min_{i \neq 1} [(1 - \lambda)\alpha_i + \lambda\beta_i] \tag{11}$$

where we have set $\lambda = \lambda_2^2$. The situation is illustrated in Fig. 2. Here we have taken $\alpha_2 \leqq \alpha_3 \leqq \cdots \leqq \alpha_g$ which we can do without loss of generality since this is merely a matter of giving names to the group elements. The bracketed expression on the right of equation (11) is plotted as the line segment with ordinate α_i at $\lambda = 0$ and ordinate β_i at $\lambda = 1$. We seek the highest point on the bottom boundary of this collection of lines, point P in Fig. 2.

Now any of the vectors $\mathbf{Y}_i$, $i = 1, 2, \cdots, g$, not just $\mathbf{Y}_1$, would serve to generate the **Y** code. We can seek a further maximization of the nearest neighbor distance (11) for the **Z** code by choice of the vector from the **Y** code to be called $\mathbf{Y}_1$. Stated otherwise, for the initial vector of the **Z** code we choose a particular vector $\mathbf{X}_1$ from the **X** code and to this we can add (directly) any of the vectors of the **Y** code. Now replacing $\mathbf{Y}_1$ by $\mathbf{Y}_l$ merely amounts to permuting the subscripts on the β_i of Fig. 2. The subscript i is replaced by k where $A_lA_i = A_k$. To combine the two codes to get the largest nearest neighbor distance, we must further

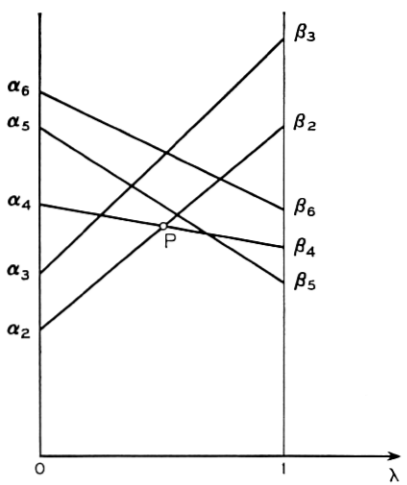

Fig. 2 — Maximum nearest neighbor distance.

maximize (11) over permutations of the β's corresponding to left translations of the group.

The maximization just considered was with respect to the manner of combining the two summand codes. There remains the matter of choosing $\mathbf{X}_1$ and $\mathbf{Y}_1$ to further increase (11). At first it might be thought that these vectors should be chosen to maximize the nearest neighbor distance in each of the summand codes. That this is not necessarily so can be seen from Fig. 2. Choosing $\mathbf{X}_1$ to increase the nearest neighbor distance in the code generated by D^1 would cause α_2 to increase. The line connecting α_2 and β_2 on the figure would move up. However, this change of $\mathbf{X}_1$ might cause α_4 to decrease by a larger amount so that point P on the figure moves downward. The situation is complicated.

The relationship between the maximum likelihood region for the sum code and the corresponding regions for the constituent codes is even more complicated in general. Let $\mathfrak{R}$ be the region belonging to $\mathbf{Z}_1$ of equation (10) and let $\mathfrak{R}^1$ and $\mathfrak{R}^2$ be corresponding regions for $\mathbf{X}_1$ and $\mathbf{Y}_1$ in the summand codes. We write $\mathbf{Z} = \lambda_1\mathbf{X} + \lambda_2\mathbf{Y}$ for a general point in the space of the direct sum representation where $\mathbf{X}$ and $\mathbf{Y}$ lie in the respective invariant subspaces of the summand codes. A point will lie in $\mathfrak{R}$ then if $|\mathbf{Z} - \mathbf{Z}_1| \leqq |\mathbf{Z} - \mathbf{Z}_i|$ for $i = 2, 3, \cdots, g$, or what is the same, if

$$\lambda_1^2 d^2(\mathbf{X}, \mathbf{X}_1) + \lambda_2^2 d^2(\mathbf{Y}, \mathbf{Y}_1) \leqq \lambda_1^2 d^2(\mathbf{X}, \mathbf{X}_i) + \lambda_2^2 d^2(\mathbf{Y}, \mathbf{Y}_i)$$

for $i = 2, 3, \cdots, g$. Thus if $\mathbf{X} \in \mathfrak{R}^1$ and $\mathbf{Y} \in \mathfrak{R}^2$ then $\mathbf{Z} \in \mathfrak{R}$, but the converse is not necessarily so in general.

A special case in which the converse holds is the following. It may happen that both the $\mathbf{X}$ code and the $\mathbf{Y}$ code have fewer than g distinct vectors. In the direct sum code (10) it may happen that each distinct vector of the $\mathbf{Y}$ code is paired at least once with each distinct vector of the $\mathbf{X}$ code. ($\mathcal{G}$ must be homeomorphic to the direct product of two groups.) In this case $\mathfrak{R}$ is the cartesian product of the two regions $\mathfrak{R}^1$ and $\mathfrak{R}^2$. The probability of no error for the sum code is given by $Q_e = Q_e^1(\lambda_1)Q_e^2(\lambda_2)$ where the factors are the probabilities of no error for the separate scaled summand codes. The information rate (3) for the sum code in this case is the weighted sum of the rates for the constituent codes

$$R = \frac{l_1}{l} R_1 + \frac{l_2}{l} R_2 .$$

We are better off using the code with the larger rate uncombined.

VI. THE CONFIGURATION MATRIX

Let $\mathbf{X}_1, \mathbf{X}_2, \cdots, \mathbf{X}_M$ be a collection of unit vectors spanning $\mathcal{S}_n$. The configuration matrix of this code is the M by M matrix ρ whose elements $\rho_{ij} = \mathbf{X}_i \cdot \mathbf{X}_j$ are the cosines of the angles between the words. Equivalent codes have identical configuration matrices except for a possible relabeling of rows and columns. This configuration matrix is real, symmetric, non-negative definite and of rank n. The diagonal elements are unity and the off-diagonal elements are of magnitude no greater than unity.

Conversely, we have the following

Lemma: Every real, symmetric, M by M non-negative definite matrix of rank n with diagonal elements unity and off-diagonal elements of magnitude less than unity is the configuration matrix of a code of M unit vectors that span $\mathcal{S}_n$.

The proof of this lemma follows readily from the fact that a real symmetric M by M matrix ρ can be diagonalized by an orthogonal matrix O, that is, $O\rho O^{-1} = \Lambda$ where, since ρ is non-negative definite and of rank n, Λ has n positive diagonal elements and all other elements are zero. Without loss of generality we can take the first n diagonal elements of Λ, say $\lambda_{ii} = \lambda_i$, $i = 1, 2, \cdots, n$ to be the positive ones. From $\rho = O^{-1}\Lambda O$, it follows that

$$\rho_{ij} = \sum_{\mu} O_{\mu i} \sqrt{\lambda_\mu}\, O_{\mu j} \sqrt{\lambda_\mu} = \mathbf{X}_i \cdot \mathbf{X}_j$$

where $\mathbf{X}_i$ is a vector of n components, the μth component being $\sqrt{\lambda_\mu} O_{\mu i}$, $i = 1, 2, \cdots, M$. We have now exhibited M unit n-vectors whose configuration matrix is the given matrix ρ. We need now only show that they span $\mathcal{S}_n$. But we have written $\rho = \tilde{X}X$ where X is the matrix of M columns and n rows whose ith column is $\mathbf{X}_i$. The tilde denotes transpose. Since the rank of a product of matrices is not greater than the smaller of the ranks of the factors, it follows that X must be of rank n, for if it were of rank less than n, so also would be ρ contrary to hypothesis. The $\mathbf{X}_i$ therefore span $\mathcal{S}_n$.

For group codes, the rows of the configuration matrix are all permutations of the first row of the matrix as can be seen from Proposition 1. Indeed the structure of this matrix is closely related to the multiplication table of the group generating the code. Let the code vector $\mathbf{X}_i = D(A_i)\mathbf{X}$, $i = 1, 2, \cdots, M$ be generated by an orthogonal representation D of a group $\mathcal{G}$ with elements $A_1, A_2, \cdots, A_M$. Here A_1 is the identity and the code need not have M distinct vectors. Denote by $\theta(A_j)$ the

angle between $\mathbf{X}_i$ and $\mathbf{X}_1$. Then $\theta(A_i^{-1}) = \theta(A_i)$, $j = 1, 2, \cdots, M$ and the configuration matrix of the code is found to be given by

$$\rho_{ij} = \cos \theta(A_i^{-1}A_j)$$

$i, j = 1, 2, \cdots, M$. If $\rho_{1j} < 1$ for $j > 1$, then the code has M distinct vectors: if $1 = \rho_{1j_1} = \rho_{1j_2} = \cdots = \rho_{1j_h}$ with $1 = j_1 < j_2 \cdots < j_h$ and these are the only elements of value unity in the first row, then the code has M/h distinct vectors.

Conversely, we have

Theorem 1: *Let $x(A_j)$ be a real-valued function defined on the elements $A_1, A_2, \cdots, A_M$ of a group $\mathcal{G}$ of order M. Let $x(A_1) = 1$, where A_1 is the identity of the group, and let $x(A_j) = x(A_j^{-1})$, $j = 1, 2, \cdots, M$. If the M by M matrix ρ with elements $\rho_{ij} = x(A_i^{-1}A_j)$ is non-negative definite and of rank n, then there exists a group code $\{M' n\}$ generated by an n-dimensional orthogonal representation of $\mathcal{G}$ that has configuration matrix ρ. Here $M' = M/h$ where h is the number of different values of j for which $x(A_j) = 1$.*

Proof: The proof follows easily from the lemma. We can find M unit vectors $\mathbf{X}_i$ (not necessarily distinct) that span $\mathcal{S}_n$ such that $\rho_{ij} = \mathbf{X}_i \cdot \mathbf{X}_j$. Without loss of generality we can suppose that $\mathbf{X}_1, \mathbf{X}_2, \cdots, \mathbf{X}_n$ are linearly independent. Now an n by n real matrix is determined by specifying its effect on n vectors that span its carrier space. For each $\mu = 1, 2, \cdots, M$ we determine the n by n matrix $D(A_\mu)$ by specifying its effect on $\mathbf{X}_1, \cdots, \mathbf{X}_n$, namely that $D(A_\mu)\mathbf{X}_i = \mathbf{X}_{l(i,\mu)}$, $i = 1, 2, \cdots, n$ where $A_\mu A_i = A_{l(i,\mu)}$. Now $\mathbf{X}_i \cdot \mathbf{X}_j = \rho_{ij} = x(A_i^{-1}A_j) = x(A_i^{-1}A_\mu^{-1}A_\mu A_j) = x(A_{l(i,\mu)}^{-1}A_{l(j,\mu)}) = \mathbf{X}_{l(i,\mu)} \cdot \mathbf{X}_{l(j,\mu)}$, so that $D(A_\mu)$ preserves the angles between n vectors spanning its carrier space. It is easy then to show that $D(A_\mu)$ preserves the angle between any two vectors and hence is an orthogonal matrix. For $j > n$,

$$\mathbf{X}_j = \sum_{h=1}^{n} \alpha_{jh}\mathbf{X}_h$$

for some set of α's. Using this representation and the orthogonality of $D(A_\mu)$, it is now easy to show that $D(A_\mu)\mathbf{X}_i = \mathbf{X}_{l(i,\mu)}$ for $i = 1, 2, \cdots, M$. The fact that D is a representation then follows readily.

Theorem 1 permits an interesting reformulation of the problem of finding an $\{M, n\}$ of largest nearest neighbor distance generated by a representation of $\mathcal{G}$. Form the modified multiplication table of $\mathcal{G}$,—an M by M array of group elements with $A_i^{-1}A_j$ in the ith row and jth column. From this table we construct a symmetric M by M matrix ρ

by replacing the group identity A_0, say, by unity, by replacing both A_1 and A_1^{-1} by the variable x_1, A_2 and A_2^{-1} by x_2, and so on. If $\mathcal{G}$ has exactly m self-reciprocal elements, there will be $K = m - 1 + (g - m)/2$ variables in ρ. The condition that ρ be non-negative definite and of rank not greater than n obtained by conditioning certain minors of ρ gives rise to polynomial constraints in the variables $x_1, \cdots, x_K$. To find the code of largest nearest neighbor distance, we must minimize $\max_i x_i$ subject to these constraints.

We notice in closing this section that the configuration matrix of an $\{M, n\}$ generated by a group $\mathcal{G}$ of order M commutes with all the matrices of the regular representation of $\mathcal{G}$ (see the appendix). Using Schur's lemma, one can then arrive at a canonical representation for configuration matrices that involves the irreducible representations of $\mathcal{G}$. But we do not pursue this topic further here.

VII. SOME THEOREMS ON DISTANCES

We now adopt the notation of the appendix. Let $\mathcal{G}$ be an abstract group of order g with elements $E, A, B, \cdots$ where E is the identity. Let $D(E), D(A), \cdots$ be a real-irreducible representation of $\mathcal{G}$ by $n \times n$ (real) orthogonal matrices. From an initial unit vector $\mathbf{X} = \mathbf{X}_E$ the representation generates a code by $\mathbf{X}_R = D(R)\mathbf{X}_E$, R runs through $\mathcal{G}$. We denote the squared distance from $\mathbf{X}_R$ to $\mathbf{X}_S$ by $d^2(\mathbf{X}_R, \mathbf{X}_S)$. We have then

$$d^2(\mathbf{X}_A, \mathbf{X}) = 2 - 2 \sum_{i,j=1}^{n} D(A)_{ij} x_i x_j$$

$$= d^2(\mathbf{X}_{RA}, \mathbf{X}_R) \tag{12}$$

for every R and $A \in \mathcal{G}$. Here $x_1, x_2, \cdots, x_n$ are the components of $\mathbf{X}$.

For codes generated from real-irreducible representations in this manner, a number of interesting distance sums are independent of the choice of the initial vector $\mathbf{X}$.

Theorem 2: *Let $D(R)$ be the matrices of a real-irreducible orthogonal representation of a group of order g. Let $\mathbf{X}_R = D(R)\mathbf{X}$. If $D(R)$ is not the trivial one-dimensional representation $D(R) = 1$, then*

$$\sum_{R \in \mathcal{G}} d^2(\mathbf{X}_R, \mathbf{X}) = 2g$$

independent of the unit vector $\mathbf{X}$.

This is really a special case of the more general

Theorem 3: *For any code generated from the initial unit vector* $\mathbf{X}$ *by a real-irreducible orthogonal representation* D *of* $\mathcal{G}$,

$$\sum_{R \in \mathcal{G}} d^2(\mathbf{X}_{R^m}, \mathbf{X}) = 2g(1 - \mu_m)$$

where

$$\mu_m = \frac{1}{gn} \sum_{R \in \mathcal{G}} \chi(R^m)$$

is a constant independent of $\mathbf{X}$. *Here* $\chi(R) = \operatorname{Tr} D(R)$ *is the character of* R *in the representation.*

Proof: Consider the matrix

$$A \equiv \sum_{R \in \mathcal{G}} D(R^m) = \sum_{R \in \mathcal{G}} D(R)D(R) \cdots D(R)$$

where there are m factors in the summand. Since the representation is by orthogonal matrices, $\tilde{D}(R) = D^{-1}(R) = D(R^{-1})$ where the tilde denotes transpose. Thus

$$\tilde{A} = \sum_{R \in \mathcal{G}} D(R^{-1}) \cdots D(R^{-1}) = A$$

since as R runs through $\mathcal{G}$ so does R^{-1}. The matrix A is thus symmetric.

We next show that A commutes with all the matrices $D(R)$. By a theorem quoted in the appendix we can then conclude that $A = \alpha I$ where I is the unit matrix. To see that A commutes with $D(R)$, consider

$$AD(R) = \sum_{S \in \mathcal{G}} D(S)^{m-1} D(S) D(R) = \sum_{S \in \mathcal{G}} D(S)^{m-1} D(SR).$$

Now set $SR = T$ so that $S = TR^{-1}$. Then

$$\begin{aligned} AD(R) &= \sum_{T \in \mathcal{G}} D(TR^{-1})^{m-1} D(T) \\ &= \sum D(TR^{-1})D(TR^{-1}) \cdots D(TR^{-1})D(T) \\ &= \sum_{T \in \mathcal{G}} D(T)D(R^{-1}T)D(R^{-1}T) \cdots D(R^{-1}T) \\ &= \sum_{U \in \mathcal{G}} D(RU)D(U)^{m-1} = D(R) \sum_{U \in \mathcal{G}} D(U)^m = D(R)A \end{aligned}$$

where we have used the substitution $U = R^{-1}T$.

From equation (12) we have

$$\begin{aligned} \sum_{R \in \mathcal{G}} d^2(\mathbf{X}_{R^m}, \mathbf{X}) &= 2g - 2 \sum_{i,j=1}^{n} x_i x_j \sum_{R \in \mathcal{G}} D(R^m)_{ij} \\ &= 2g - 2 \sum_{i,j=1}^{n} x_i x_j A_{ij} = 2(g - \alpha) \end{aligned}$$

by the diagonal property of A just established. To find α consider the trace of A. We have

$$\text{Tr } A = \text{Tr } \alpha I = \alpha n = \sum_{R \in \mathcal{G}} \text{Tr } D(R^m) = \sum_{R \in \mathcal{G}} \chi(R^m).$$

The theorem then follows.

To establish Theorem 2, notice that for the trivial representation $D^1(R) = 1$, we have $\chi^1(R) = 1$. For the character $\chi(R)$ of any other nonequivalent real-irreducible representation we then have

$$\sum_{R \in \mathcal{G}} \chi^1(R)\chi(R) = \sum_{R \in \mathcal{G}} \chi(R) = 0$$

by the orthogonality relations (appendix). Using this fact and setting $m = 1$ in Theorem 3 yields Theorem 2.

Theorem 4: *Let* $\mathcal{C}$ *be a class of* n_c *elements of* $\mathcal{G}$ *with character* $\chi(\mathcal{C})$. *For any code generated from the initial unit vector* $\mathbf{X}$ *by a real-irreducible orthogonal representation* D *of* $\mathcal{G}$,

$$\sum_{R \in \mathcal{C}} d^2(\mathbf{X}_R , \mathbf{X}) = 2n_c\left(1 - \frac{1}{n}\chi(\mathcal{C})\right) \tag{13}$$

independent of the unit vector $\mathbf{X}$.

Proof:

$$\sum_{R \in \mathcal{C}} d^2(\mathbf{X}_R , \mathbf{X}) = 2n_c - 2 \sum x_i x_j \sum_{R \in \mathcal{C}} D(R)_{ij} . \tag{14}$$

Now consider the matrix

$$B = \sum_{R \in \mathcal{C}} D^\alpha(R) = \frac{n_c}{g} \sum_{S \in \mathcal{G}} D^\alpha(SRS^{-1}) = \frac{n_c}{g} \sum_{S \in \mathcal{G}} D^\alpha(S)D^\alpha(R)D^\alpha(S^{-1})$$

where $D^\alpha(R)$ is an irreducible (over the complex field) representation of dimension m of $\mathcal{G}$. Now B commutes with all the matrices of D^α since

$$BD^\alpha(T) = \frac{n_c}{g} \sum_{S \in \mathcal{G}} D^\alpha(S)D^\alpha(R)D^\alpha(S^{-1}T)$$

$$= \frac{n_c}{g} \sum_{U \in \mathcal{G}} D^\alpha(TU^{-1})D^\alpha(R)D^\alpha(U) = D^\alpha(T)B$$

where we have set $S^{-1}T = U$. By Schur's lemma, $B = kI$ where I is the m by m unit matrix. Taking traces we have

$$\text{Tr } B = \text{Tr} \sum_{R \in \mathcal{C}} D^\alpha(R) = n_c\chi^\alpha(\mathcal{C}) = \text{Tr } kI = km$$

so that

$$B = \sum_{R \in \mathcal{C}} D^{\alpha}(R) = \frac{n_c}{m} \chi^{\alpha}(\mathcal{C}) I. \tag{15}$$

If now the real-irreducible orthogonal representation D is also irreducible, by equation (15) the inner sum in (14) is $(n_c/n)\chi(\mathcal{C})\delta_{ij}$ and the theorem (13) follows at once.

Suppose now that D is not irreducible. Then (see appendix) D is equivalent to an orthogonal representation of the form

$$D'(R) = \begin{pmatrix} U^{\alpha}(R) & V^{\alpha}(R) \\ -V^{\alpha}(R) & U^{\alpha}(R) \end{pmatrix} \tag{16}$$

where $D^{\alpha}(R) = U^{\alpha}(R) + iV^{\alpha}(R)$ is an irreducible representation by unitary matrices and U^{α} and V^{α} are real and of dimension m where $n = 2m$. We can suppose the D of equation (14) to be of the form (16). Now let

$$\hat{B} = \sum_{R \in \mathcal{C}} D'(R)$$

and set

$$U = \frac{1}{\sqrt{2}} \begin{pmatrix} I & iI \\ iI & I \end{pmatrix}$$

where as before I is the m by m unit matrix. One then finds by direct computation that

$$U^{-1}\hat{B}U = \sum_{R \in \mathcal{C}} \begin{pmatrix} D^{\alpha}(R) & 0 \\ 0 & D^{\alpha}(R)^* \end{pmatrix}$$

$$= \frac{n_c}{m} \begin{pmatrix} \chi^{\alpha}(\mathcal{C}) I & 0 \\ 0 & \chi^{\alpha}(\mathcal{C})^* I \end{pmatrix} \equiv H$$

where the middle equality follows from equation (15). Now let $\chi^{\alpha}(\mathcal{C}) = \mu + i\nu$ with μ and ν real. Direct computation gives

$$\hat{B} = UHU^{-1} = \frac{n_c}{m} \begin{pmatrix} \mu I & \nu I \\ -\nu I & \mu I \end{pmatrix}. \tag{17}$$

The right of (14) is

$$2n_c - 2 \sum \hat{B}_{ij} x_i x_j$$

and using (17) this becomes

$$2n_c - \frac{2n_c}{m}\mu \sum x_i^2 = 2n_c\left(1 - \frac{\mu}{m}\right).$$

From equation (16), however, $\chi(\mathcal{C}) = 2\mu$, so that (13) then follows and the theorem is proved in all cases.

Since every group code can be thought of as the direct sum of codes generated by real-irreducible representations, and since squared distance in the sum code is the sum of squared distances in the separate codes, Theorems 2, 3, and 4 have ready analogues for all group codes. For example, if a group code does not contain the identity representation, then

$$\sum_{R \in \mathcal{G}} d^2(\mathbf{X}_R , \mathbf{X}) = 2g.$$

Theorems 3 and 4 hold for group codes in general when $\chi(R)$ is replaced by the weighted sum $\sum \lambda_i^2 \chi^i(R)$ of the characters of the constituent real-irreducible codes.

Another theorem of interest concerning codes generated from any group of orthogonal matrices arises from the fact (12) that $d^2(\mathbf{X}_A , \mathbf{X}) = d^2(\mathbf{X}_{RA} , \mathbf{X}_R)$. Let there be a point of the code distant d from the point $\mathbf{X}_E$. Starting from $\mathbf{X}_E$, we imagine moving from word to word of the code restricting our moves so that from any word we can move only to a word distant d away. We shall call the collection of words that can be reached from $\mathbf{X}_E$ in this manner "a d chain starting from $\mathbf{X}_E$". $\mathbf{X}_E$ is to be included in this chain.

Theorem 5: *Let the words of a d chain starting from* $\mathbf{X}_E$ *be* $\mathbf{X}_E$, $\mathbf{X}_{A_1}$, $\mathbf{X}_{A_2}$, $\cdots$ $\mathbf{X}_{A_h}$. *Then the group elements* E, A_1, A_2, $\cdots$, A_h *form a subgroup* $\mathfrak{H}$ *of* $\mathcal{G}$. *The group elements whose corresponding words are distant d from* $\mathbf{X}_E$ *form a set of generators for* $\mathfrak{H}$. *If* $\mathfrak{H}$ *is a proper subgroup of* $\mathcal{G}$, *then from any word corresponding to a group element not in* $\mathfrak{H}$, *a new d chain may be formed and the group elements corresponding to the points of this new d chain will form a coset of* $\mathfrak{H}$.

Proof: Suppose all the points distant d from $\mathbf{X}_E$ are $\mathbf{X}_{A_1}$, $\mathbf{X}_{A_2}$, $\cdots$, $\mathbf{X}_{A_m}$. Let us construct a table of group elements in the following manner. The first row is E, A_1, A_2, $\cdots$, A_m. The $K + 1$st row of the table is formed from the preceding K rows as follows. We examine the elements of the table in order, reading from left to right in the first row, then from left to right in the second row, and so on. Let R be the first element arrived at in reading the first K rows that does not appear in the first

column, rows 1, 2, $\cdots$, K. The the $K+1$st row is to be

$$R, RA_1, RA_2, \cdots, RA_m.$$

The table thus appears

$$\begin{array}{llll} E, & A_1, & A_2, & A_m \\ A_1, & A_1^2, & A_1A_2, & A_1A_m \\ A_2, & A_2A_1, & A_2, & A_2A_m \\ \vdots & & & \\ A_m, & A_mA_1, & A_mA_2, & A_m^2 \\ B, & BA_1, & BA_2, & BA_m \\ \vdots & & & \\ R, & RA_1, & RA_2, & RA_m \end{array}$$

When j rows have been written and every element in these j rows has appeared once in the first column the process is stopped and the table is considered complete. The table can have at most g rows. Now from $d^2(\mathbf{X}_R, \mathbf{X}_E) = d^2(\mathbf{X}_{SR}, \mathbf{X}_S)$, it follows that the words represented by the elements in the 2nd, 3rd, $\cdots$, $m+1$st columns of the Kth row are all distant d from the word represented by the element in the first column of the Kth row. Furthermore, these m words are all the words of the code that are distant d from the word represented by the element in the first column of the Kth row. Thus the elements of the first column of the table give the points of the d chain starting from $\mathbf{X}_E$. That these elements of the first column form a group $\mathfrak{IC}$ and that $A_1, A_2, \cdots, A_m$ are generators of $\mathfrak{IC}$ is clear from the method of constructing the table, for we have formed all possible distinct products of the A's and listed the distinct elements thus obtained in the first column. Let $\mathfrak{IC}$ be a proper subgroup of $\mathfrak{G}$ and let S be an element of $\mathfrak{G}$ not in $\mathfrak{IC}$. If we multiply every element in the above table by S, we obtain a new table giving all the points that can be reached from point $\mathbf{X}_S$ by steps of distance d. The first column of this table lists the points of the d chain starting from $\mathbf{X}_S$ and the corresponding elements are just the coset $S\mathfrak{IC}$ of $\mathfrak{IC}$.

VIII. BINARY GROUP CODES

The group codes (n, k) for the binary channel introduced in Ref. 21 are group codes in the present sense. Each word of an (n, k) code is an

n-place binary sequence. Replace each zero by 1 and replace each 1 by -1 in each word. Then write each word (a sequence of ± 1's) as a diagonal $n \times n$ matrix. This collection of 2^k $n \times n$ orthogonal matrices forms an Abelian group $\mathfrak{B}_k$ that is isomorphic to the k-fold direct product of the simple two element Abelian group. The matrices generate the code by operating on the n-vector $(1, 1, 1, \cdots, 1)$. The real-irreducible representations of this group are all one dimensional. There are 2^k of them. The representation by $n \times n$ matrices just considered is already exhibited in reduced form as the direct sum of n of these real-irreducible representations.

IX. CONCLUDING REMARKS

The foregoing paragraphs outline some of the theory of group codes for the Gaussian channel. The development of this subject is clearly incomplete: we have raised more questions than we have answered. Perhaps the outstanding problem is that of finding a tractable method of choosing the initial vector to maximize the nearest neighbor distance.

There is a great abundance of groups of arbitrarily large order that can be examined from the point of generating group codes. The symmetric group and the hyperoctahedral group appear most promising for initial investigation since their structure and irreducible representations (which are all real) are comparatively well understood.

APPENDIX

Review of Group Representation Theory[25]

Let $\mathcal{G}$ be a finite group of order g with elements $E, A, B, \cdots$. The letters R and S will be used for the general element of $\mathcal{G}$ and E will denote the identity of $\mathcal{G}$. As R runs through $\mathcal{G}$, the distinct elements of the set RAR^{-1} are said to form a class of $\mathcal{G}$. The elements A and B are said to belong to the same class of $\mathcal{G}$ if there exists an S such that $A = SBS^{-1}$. $\mathcal{G}$ can be divided uniquely into a union of classes, no two classes containing a common element. The number of elements in a class of $\mathcal{G}$ is a divisor of g.

If $\mathcal{H}$ is a subgroup of $\mathcal{G}$ and if $\mathcal{H}$ is of order h, then h is a divisor of g and the number g/h is called the index of $\mathcal{H}$ under $\mathcal{G}$. If, for every R in $\mathcal{H}$, all elements of $\mathcal{G}$ in the same class as R are also contained in $\mathcal{H}$, then $\mathcal{H}$ is said to be a self-conjugate subgroup of $\mathcal{G}$. A subgroup $\mathcal{H}$ of $\mathcal{G}$ is said to be proper if $h < g$.

The matrices in what follows are assumed to have elements in the field of complex numbers.

If to every element R of a finite group $\mathcal{G}$ there corresponds an n by n nonsingular matrix $D(R)$ and if $D(R)D(S) = D(RS)$, the collection of matrices $\Delta = \{D(R), R \text{ runs through } \mathcal{G}\}$ is said to form an n-dimensional representation of $\mathcal{G}$. The matrices of Δ form a group under matrix multiplication. If the correspondence between the matrices of Δ and the elements of $\mathcal{G}$ is one-to-one, Δ is said to be a faithful representation of $\mathcal{G}$. If for some $R \neq S$, $D(R) = D(S)$, Δ is said to be an unfaithful representation of $\mathcal{G}$. The matrix $D(E)$ is always the n by n unit matrix. If a representation is unfaithful, the elements represented by $D(E)$ form a self-conjugate subgroup of $\mathcal{G}$, say of order h, and to each matrix of Δ correspond exactly h elements of $\mathcal{G}$. Δ contains g/h distinct matrices. If $D(E)$, $D(A)$, $\cdots$ is an n dimensional representation of $\mathcal{G}$, so is $MD(E)M^{-1}$, $MD(A)M^{-1}$, $\cdots$ where M is any nonsingular n by n matrix. The two representations Δ and $M\Delta M^{-1}$ are called equivalent. Every representation of a finite group is equivalent to a representation by unitary matrices. Henceforth we shall be concerned only with such unitary representations.

A finite collection of n by n matrices $O_1, O_2, \cdots, O_K$ is said to be reducible if there exists an n by n unitary matrix U such that for $i = 1, 2, \cdots, K$ we have

$$UO_iU^{-1} = \left| \begin{array}{c|c} A_i & D \\ \hline C & B_i \end{array} \right|$$

where A_i is an l by l matrix, B_i is an n–l by n–l matrix, $0 < l < n$, C is an n–l by l matrix all of whose elements are zero, and D is an l by n–l matrix all of whose elements are zero. It is assumed that l is independent of i. A collection of matrices that is not reducible is said to be irreducible.

Every finite group has exactly as many nonequivalent irreducible representations as it has classes. If $l_1, l_2, \cdots, l_c$ are the dimensions of all the nonequivalent irreducible representations of $\mathcal{G}$, of order g, then

$$\sum_1^c l_i^2 = g.$$

If $D^\alpha(R)_{\mu\nu}$ is the element in the μth row and νth column of the matrix representing R in the l_α-dimensional irreducible representation, α, of $\mathcal{G}$, then

$$\sum_{R \in \mathcal{G}} D^\alpha(R)_{\mu\nu} D^\alpha(R)^*_{\mu',\nu'} = \delta_{\mu\mu'}\,\delta_{\nu\nu'}\,g/l_\alpha \qquad \mu, \mu', \nu, \nu' = 1, 2, \cdots, l_\alpha .$$

Here * means complex conjugate and δ is the usual Kronecker symbol. If the matrices $D^\beta(R)$ form an l_β dimensional irreducible representation

of $\mathcal{G}$ not equivalent to the representation α, then

$$\sum_{R \in \mathcal{G}} D^{\alpha}(R)_{\mu\nu} D^{\beta}(R)^{*}_{\mu'\nu'} = 0,$$

$$\mu, \nu = 1, 2, \cdots, l_{\alpha}, \quad \mu', \nu' = 1, 2, \cdots, l_{\beta}.$$

If $D(R)$ is the n by n matrix representing R in the representation Δ, the trace of $D(R)$, namely

$$\chi(R) = \sum_{\mu=1}^{n} D(R)_{\mu\mu},$$

is called the character of R in the representation Δ. If R and S are in the same class of $\mathcal{G}$, then $\chi(R) = \chi(S)$, for any representation of $\mathcal{G}$. The characters of the irreducible representations α and β of $\mathcal{G}$ satisfy the orthogonality conditions

$$\sum_{R \in \mathcal{G}} \chi^{\alpha}(R)\chi^{\beta}(R)^{*} = g\, \delta_{\alpha\beta}.$$

Here $\delta_{\alpha\beta}$ is unity if α and β are equivalent representations and is zero otherwise.

Let Δ be any representation of $\mathcal{G}$ with character $\chi(R)$. Let the characters of the irreducible representations of $\mathcal{G}$ be $\chi^{j}(R)$, $j = 1, 2, \cdots, c$ where c is the number of nonequivalent irreducible representations of $\mathcal{G}$ (= number of classes of $\mathcal{G}$). Then $\chi(R)$ may be written uniquely in the form

$$\chi(R) = \sum_{j=1}^{c} a_j \chi^{j}(R), \qquad \text{all } R \text{ in } \mathcal{G},$$

where the a_j are nonnegative integers independent of R. In fact,

$$a_j = \frac{1}{g} \sum_{R \in \mathcal{G}} \chi(R)\chi^{j}(R)^{*}.$$

The representation Δ is said to contain the irreducible representation j a_j times and there exists a unitary matrix U independent of R such that

$$UD(R)U^{-1} = \begin{vmatrix} D^{i}(R) & 0 \cdots 0 \\ 0 & D^{j}(R) \cdots 0 \\ \vdots & \vdots \quad \vdots \\ 0 & 0 \quad D^{k}(R) \end{vmatrix} \qquad \text{all } R \text{ in } \mathcal{G}$$

where $D^i(R)$, $D^j(R)$, $D^k(R)$ etc., are matrices of the ith, jth, kth, etc., irreducible representation of $\mathcal{G}$, and 0 stands for the appropriate matrix with all elements zero. The jth irreducible representation will occur exactly a_j times among $D^i(R)$, $D^j(R)$, $D^k(R)$ and so on.

Every group $\mathcal{G}$ possesses a faithful representation, called the regular representation Γ, that consists of g by g permutation matrices. The rows and columns of these matrices can be labelled by the elements of $\mathcal{G}$. The entry in row R and column S of the matrix representing T is unity if $R = TS$ and is zero otherwise. The regular representation is reducible: it contains the irreducible representation D^α exactly l_α times, $\alpha = 1, 2, \cdots, c$.

Let $D'(R)$ and $D''(R)$, R runs through $\mathcal{G}$, be irreducible representations of $\mathcal{G}$ of dimension d' and d'' respectively. Let the matrix H satisfy $D'(R)H = HD''(R)$ for all R in $\mathcal{G}$. Then either H is the zero matrix or H is square and nonsingular so that $d' = d''$, and the two representations are equivalent. A matrix that commutes with all the matrices of an irreducible representation of $\mathcal{G}$ is a multiple of the unit matrix. These statements are known as Schur's lemma.

Much of the foregoing remains valid with minor modifications when the number field in question is the real rather than the complex numbers. One easily finds that every real representation of a finite group is equivalent (over the reals) to a representation by orthogonal matrices. The only real symmetric matrix that commutes with all the matrices of a real-irreducible representation is a multiple of the unit matrix. If $D^\alpha(R)$ and $D^\beta(R)$ are nonequivalent real-irreducible representations by real orthogonal matrices, respectively of dimension l_α and l_β, then

$$\sum_{R \in \mathcal{G}} D^\alpha(R)_{\mu\nu} D^\beta(R)_{\mu'\nu'} = 0,$$

$$\mu, \nu = 1, 2, \cdots, l_\alpha \quad \mu', \nu' = 1, 2, \cdots, l_\beta,$$

$$\sum_{R \in \mathcal{G}} [D^\alpha(R)_{\mu\nu} D^\alpha(R)_{\mu'\nu'} + D^\alpha(R)_{\mu\nu'} D^\alpha(R)_{\mu'\nu}] = 2\,\delta_{\mu\mu'}\,\delta_{\nu\nu'} g/l_\alpha,$$

$$\mu, \nu, \mu', \nu' = 1, 2, \cdots l_\alpha.$$

For the characters one has

$$\sum_{R \in \mathcal{G}} \chi^\alpha(R)\chi^\beta(R) = 0$$

if the representations α and β are not equivalent, while

$$\sum_R [\chi^\alpha(R)\chi^\alpha(R) + \chi^\alpha(R^2)] = 2g.$$

Every real-irreducible representation that is reducible (over the complex numbers) is equivalent to the direct sum of an irreducible representation and its complex conjugate. If the irreducible unitary representation $D(R) = U(R) + iV(R)$, with U and V real, is not equivalent to a real orthogonal representation, then

$$\left(\begin{array}{c|c} U(R) & V(R) \\ \hline -V(R) & U(R) \end{array}\right) \tag{18}$$

is a real-irreducible representation by real orthogonal matrices.

For an irreducible representation $D(R)$ with character $\chi(R)$, the sum

$$h = \frac{1}{g} \sum_{R \in \mathcal{G}} \chi(R^2)$$

can have only one of the three different values $0, \pm 1$. If $h = 1$, $D(R)$ is equivalent to a representation by real orthogonal matrices. If $h = -1$, the representation D is equivalent to its complex conjugate, but is not equivalent to a real representation. A real-irreducible representation can be made from each irreducible representation D having $h = -1$ by forming real matrices of the form (18), where U and V are the real and imaginary parts of D. Finally, if $h = 0$, D is not equivalent to its complex conjugate and is not equivalent to a real representation. Nonequivalent irreducible representations for which $h = 0$ occur then in complex conjugate pairs. Each such pair gives rise to a single real-irreducible representation through the recipe (18). Thus, finally, if h has the value 0 for exactly $2p$ of the c nonequivalent irreducible representations of $\mathcal{G}$, then $\mathcal{G}$ has exactly $c - p$ nonequivalent real-irreducible representations.

REFERENCES

1. Kotel'nikov, V. A., Thesis, Molotov Energy Institute, Moscow, 1947, translated as *The Theory of Optimal Noise Immunity,* New York: McGraw-Hill Book Co., 1959.
2. Shannon, C. E., "A Mathematical Theory of Communication," B.S.T.J., *27,* Nos. 3 and 4 (July and October 1948), pp. 379–423; 623–656.
3. Shannon, C. E., "Communication in the Presence of Noise," Proc. IRE, *37,* (January 1949), pp. 10–21.
4. Rice, S. O., "Communication in the Presence of Noise—Probability of Error for Two Encoding Schemes," B.S.T.J., *29,* No. 1 (January 1950), pp. 60–93.
5. Gilbert, E. N., "A Comparison of Signalling Alphabets," B.S.T.J., *31,* No. 3 (May 1952), pp. 504–522.
6. Shannon, C. E., "Probability of Error for Optimal Codes in a Gaussian Channel," B.S.T.J., *38,* No. 3 (May 1959), pp. 611–656.
7. Wolfowitz, J., *Coding Theorems of Information Theory,* Berlin: Springer-Verlag, 1961.
8. Stutt, C. A., "Information Rate in a Continuous Channel for Regular-Simplex Codes," IRE Trans. *IT-6* (December 1960), pp. 516–522.

9. Balakrishnan, A. V., "A Contribution to the Sphere-Packing Problem of Communication Theory," J. Math. Anal. Applications, *3* (December 1961), pp. 485–506. "Signal Selection Theory for Space Communication Channels," Chapter 1 in *Advances in Communication Systems,* ed. A. V. Balakrishnan, New York: Academic Press, 1965.
10. Slepian, D., "Bounds on Communication," B.S.T.J., *42,* No. 3 (May 1963), pp. 681–707.
11. Gallager, R. G., "A Simple Derivation of the Coding Theorem and Some Applications," IEEE Trans., *IT-11* (January 1965), pp. 3–18.
12. Wyner, A. D., "Capabilities of Bounded Discrepancy Decoding," B.S.T.J., *44,* No. 6 (July–August 1965), pp. 1061–1122.
13. Wozencraft, J. M. and Jacobs. I. M., *Principles of Communication Engineering,* New York: John Wiley & Sons, 1965.
14. Weber, C. L., "On Optimal Signal Selection for M-ary Alphabets with Two Degrees of Freedom," IEEE Trans., *IT-11* (April 1965), pp. 299–300, and "New Solution to the Signal Design Problem for Coherent Channels," IEEE Trans., *IT-12* (April 1966), pp. 161–167.
15. Slepian, D., "Permutation Modulation," Proc. IEEE *53* (March 1965), pp. 228–236.
16. Zetterberg, L. H., "A Class of Codes for Polyphase Signals on a Band-limited Gaussian Channel," IEEE Trans., *IT-11* (July 1965), pp. 385–395. "Detection of a Class of Coded and Phase-Modulated Signals," IEEE Trans., *IT-12* (April 1966), pp. 153–161.
17. Peterson, W. W. and Kasami, T., "Reliability Bounds for Polyphase Codes for the Gaussian Channel," Scientific Report No. 3 (July 1965), Dept. of Elec. Eng., U. of Hawaii. Abstract appears in IEEE Trans., *IT-12,* No. 2 (April 1966), p. 277.
18. Reed, I. S. and Scholtz, R. A., "N-Orthogonal Phase-Modulated Codes," IEEE Trans., *IT-12* (July 1966), pp. 388–395.
19. Scholtz, R. A. and Weber, C. L., "Signal Design for Phase-Incoherent Communications," IEEE Trans., *IT-12* (October 1966), pp. 456–463.
20. Landau, H. J. and Slepian, D., "On the Optimality of the Regular Simplex Code," B.S.T.J., *45,* No. 8 (October 1966), pp. 1247–1272.
21. Slepian, D., "A Class of Binary Signaling Alphabets, B.S.T.J., *35,* No. 1 (January 1956), pp. 203–234.
22. Dunn, James G., "Coding for Continuous Sources and Channels," Thesis, Electrical Engineering Department, Columbia University, May 1965.
23. Coxeter, H. S. M., *Regular Polytopes,* New York: The Macmillan Company, 1963.
24. Robinson, G. de B., "On the Fundamental Region of an Orthogonal Representation of a Finite Group," Proc. London Math. Soc., *43,* Sec. 2 (1937), pp. 289–301.
25. Boerner, H., *Representation of Groups,* Amsterdam: North-Holland Publishing Co., 1963.
Murnaghan, F. D., *The Theory of Group Representations,* Baltimore: Johns Hopkins Press, 1938.
Weyl, H., *The Classical Groups,* Princeton: Princeton University Press, 1946.
Wigner, E. P., *Group Theory,* New York: Academic Press, 1959.
Lomont, J. S., *Applications of Finite Groups,* New York: Academic Press, 1959.