*Mathematical Geology, Vol. 8, No. 1, 1976*

# Delineation and Analysis of Clusters in Orientation Data[1]

**R. J. Shanley[2] and M. A. Mahtab[3]**

*This paper presents a technique for (1) clustering orientation data with minimum constraint on resulting partitions, and (2) testing clusters against a probability distribution defined on the unit sphere which admits elliptical symmetry about its mean. The use of an objective function to highlight certain features of the data is discussed. The technique for delineation and analysis of clusters is applied to an example problem through use of a computer code.* KEY WORDS: cluster analysis, statistics, numerical taxonomy, structure.

## INTRODUCTION

An understanding of many phenomena in the physical and earth sciences necessitates the collection and analysis of orientation data. Examples of such data are: axes of crystals and glacial till pebbles and fracture orientations. The orientation data may be clustered around one or more modes or statistically preferred orientations. Delineation of these clusters (or sets) in the data as well as estimation of the mean direction for each cluster (along with determining the precision of this estimate) are essential for studying the various problems involving orientation data.

An example in point is the mathematical modeling of a fractured rock mass for numerical analysis of boundary-value problems encountered in subsurface excavations. If these problems are solved using a representative volume element (which is an aggregate of many fracture and intact rock elements, see Mahtab, 1974), the mean directions (and dispersions) of fracture sets are reflected in the resulting mean stresses or strains (and their dispersion). John (1968) illustrates the influence of fracture orientations on the shape and direction of movement of a wedge of rock in a roadcut; note that, near the position of limiting equilibrium, the "safety factor" is sensitive

[1] Manuscript received 7 May 1974; revised 27 July 1974; publication authorized by the Office of Mineral Information, U.S. Bureau of Mines.

[2] U.S. Army Mobility Equipment Research Division, Fort Belvoir, Virginia.

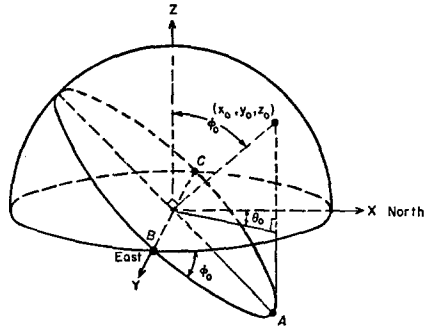
[3] Acres Consulting Services Limited, Niagara Falls, Canada; formerly with U.S. Bureau of Mines, Denver, Colorado.

to small changes (of the order of 5 degrees) in the orientations of the fracture planes delimiting the wedge. In a study of the influence of fractures on cavability (collapse) of rock mass above an "undercut" in a block-caving mine (Mahtab and Dixon, 1975), the extent of the shear failure zone was shown to increase dramatically when the dip of the fracture set permeating the rock mass was changed from 15° to 0°.

Because orientation measurements can be represented conveniently by unit vectors, it is natural that cluster-analysis techniques and probability-distribution theory be applied to collections of data points on the unit hemisphere.

Several partitions of a sample of orientations are defined using an adaptation of Wishart's (1968) one-level mode analysis. An objective function is evaluated at each partition. The clusters belonging to partitions near the point at which the function assumes its minimum are compared (using the chi-square test) with a probability distribution defined on the surface of the unit sphere, studied by Bingham (1964), that allows elliptical symmetry about the mean of a given cluster. The new technique for delineation and analysis of clusters in orientation data described herein affords increased flexibility over existing techniques (Mahtab and others, 1972). The authors have developed a computer code incorporating the new technique. The computer code was made available to potential users in a Bureau of Mines report (see Shanley and Mahtab, 1975). An application of the technique is made to an example data set composed of orientations of fractures in a porphyry copper deposit.

Axial orientation data can be represented in various manners. Given any




$\phi_0$ - Dip of plane *ABC*
$\theta_0$ - Azimuth of dip of plane
$X_0$ - Sin $\phi_0$ cos $\theta_0$
$Y_0$ - Sin $\phi_0$ sin $\theta_0$
$Z_0$ - Cos $\phi_0$

Figure 1. Representation of fracture orientations by points on unit hemisphere.

axis (for example, a normal to a fracture plane), we may construct a line in three-dimensional rectangular coordinate space which passes through the coordinate origin and is parallel to the given axis. The spherical coordinates $(1, \theta, \phi)$ of the point of intersection of the constructed line with the unit hemisphere uniquely determine the orientation (but not the position) of the axis (Fig. 1).

The area of concern to the authors is the study of rock fractures and their influence on the mechanical behavior of structures excavated in a fractured rock mass. To further this study it is necessary to extract, from sets of fracture orientation data, information on the preferred orientation and the dispersion of fracture sets, if such information exists.

In general, it is impossible to determine beforehand the distribution of orientations of fractures in a fracture set. However, the distribution usually does not have circular symmetry about the mean of the set. Elliptical symmetry about the mean may be more common.

## CLUSTERING OF ORIENTATION DATA

For the aforementioned reasons, a technique used in the delineation and analysis of clusters in orientation data (especially fracture orientations) should incorporate a clustering algorithm which is unrestrictive as to the size and shape of clusters. In addition, the algorithms used for clustering should not presuppose the distribution to be fitted to the data points in the resulting clusters. Note that the clustering technique presented in this section, and the technique for analysis of clusters, using Bingham distribution, presented in the next section, can be used independently of each other.

In what follows, a partition $P$ of a set of orientation data $A = \{(\theta_1, \phi_1), (\theta_2, \phi_2), \ldots, (\theta_n, \phi_n)\}$ will be a collection $\{A_i\}$ of distinct nonempty subsets of $A$ such that each observation in $A$ is contained in some $A_i$. The elements $A_i$ of partition $P$ will be referred to as clusters.

### Partitioning of Data Using a Modified One-Level Mode Analysis

To apply the one-level algorithm to a set of orientations $A = \{(\theta_1, \phi_1), (\theta_2, \phi_2), \ldots, (\theta_n, \phi_n\}$, two numbers must be chosen: $r$, a positive real number, and $k$, a positive integer. In addition, a "similarity" function $\rho\ (X_i, X_j)$ must be defined on the set of all pairs $(X_i = (\theta_i, \phi_i), X_j = (\theta_j, \phi_j))$ of orientations, to measure the similarity of any two axes. Spheres of radius $r$ then are constructed about every datum. The centers of those spheres containing $k$ or more observations are called dense points.

The remainder of the one-level algorithm consists of two steps: (1) partitioning of the dense points, and (2) defining a partition of all of $A$ using the nondense elements of $A$ and the partition defined on the dense points.

In step 1 we start with the partition $P_1$ (on the set of dense points) in which every cluster contains a single point. $P_1$ then is modified, to obtain a new partition $P_2$, by combining groups of clusters of $P_1$ into single clusters. The rule for determining whether two dense points, $v$ and $w$ say, will belong to the same cluster in $P_2$ is as follows: if there are dense points $X_0(=v)$, $x_1$, . . ., $X_m(=w)$, such that $\rho(X_i, X_{i+1}) \leqslant r$, $i = 0,1,2, \ldots, m-1$, then $v$ will $w$ will be in the same cluster of $P_2$.

Step 2 consists of modifying $P_2$, thereby obtaining a partition $P_3$ defined on $A$, and not on the set of dense points alone. This is done by adding nondense points to the clusters in $P_2$ until every point in $A$ belongs to some cluster. The authors assign a nondense point to the cluster containing the nearest dense point. This criterion is the choice of the authors and is not the only criterion that could be used (Wishart, 1968, p. 298). The one-level algorithm stops when $P_3$ has been defined. No dense point may be removed from its original partition element, and the number of partition elements may not change.

The strong point of the algorithm is the absence of restriction on size or shape of clusters, a restriction present, for example, in hierarchical clustering algorithms employing minimum variance as the objective function.

A drawback of the one-level algorithm, recognized by Wishart (1968), is the element of external control involved in allowing user selection of both $r$ and $k$. Furthermore, some clusters might not contain data modes, depending on the sample and choice of $r$ and $k$. These difficulties are due chiefly to the arbitrariness of the selection of $r$ and $k$. For this reason, the authors have modified the algorithm by allowing the selection of $r$ to remain arbitrary (subject only to $r \geqslant 0$) and selecting $k$ by means of a test for randomness derived from the Poisson distribution (Mahtab and others, 1972). If we allow $n$ to be sample size (number of points in $A$), and $c$ the percentage of the area of $S$ ($S = \{(x,y,z) \mid x^2+y^2+z^2 = 1, z \geqslant 0\}$) lying within a sphere of radius $r$ centered at a point of $S$, then $k$ is defined to be the smallest integer $K$ such that

$$P(D > K-1) = 1 - \sum_{j=0}^{K} e^{-m} m^j/j! \leqslant 0.05$$

where $m = cn$, and $P$ is the probability of occurrence of a random density $D$.

According to the Poisson model, $k$ is that integer such that the probability of a random grouping of $k$ or more points lying inside a sphere of radius $r$ centered at a point of $S$ is less than 5 percent. This modification not only mitigates the objectionable external control in user selection of both $r$ and $k$, but also increases the likelihood that clusters defined by the algorithm will correspond to data modes.

It remains to define the similarity measure $\rho$. Let $X_1 = (\theta_1, \phi_1)$, $X_2 = (\theta_2, \phi_2)$ be points in $A$. The coordinates $(\theta_1, \phi_1)$, $(\theta_2, \phi_2)$ give the attitudes of

two (possibly collinear) axes intersecting at the origin. We define the distance between these axes to be

$$\rho(X_1,X_2) = \alpha, \tag{1}$$

where $\alpha$ is the acute angle of intersection of the two axes. Thus $\rho$ measures the "similarity" of any two axes taken from $A$.

### Selecting Best Partition Using an Objective Function

In general, two partitions defined by the previously discussed clustering technique will be different if the corresponding values of $r$ are sufficiently different. In the event that more information than is provided by the analysis is desired by the user, an "objective function" may be defined on the collection of partitions for the purpose of determining that partition for which the function exhibits desired behavior. This usually consists of minimizing or maximizing some variable associated with the data, such as variance. The objective function offered here is

$$F(P) = \sum_{j=1}^{M} \sum_{i=1}^{N_j} d^2(X_j^i, \bar{X}_j) + \sum_{i=1}^{M-1} \sum_{j=i+1}^{M} d^2(\bar{X}_i, \bar{X}_j) \tag{2}$$

where $M$ is the number of clusters in partition $P$, $N_j$ is the number of points in cluster $j$ of partition $P$, $X_j^i$ is the $i$th element of cluster $j$, $d(X_i,X_j)$ is the Euclidean distance between $X_i$ and $X_j$, and $\bar{X}_i$ is the center of gravity of cluster $i$, assuming each data point has a unit mass.

The center of gravity of a set of orientation data $C = \{(\theta_i, \phi_i), i = 1,2, \ldots, m\}$ is given, in rectangular coordinates, by

$$\bar{X}_c = \left( \sum_{i=1}^{m} x_i/m, \sum_{i=1}^{m} y_i/m, \sum_{i=1}^{m} z_i/m \right)$$

(see Fig. 1 and Mahtab and others, 1972, p. 7–8).

The function $F(P)$ was empirically tested on generated sets of data, having circular symmetry about their means, and provided good results in that it assumed its minimum at partitions whose clusters were close to the ones generated.

## ANALYSIS OF CLUSTERS OF ORIENTATION DATA

After examining many sets of fracture orientation data, the authors have come to the conclusion that requiring the candidate distribution to exhibit circular symmetry about the mean is unrealistic. In the majority of situations examined, circular symmetry about the mean is not present. However, there is a distribution that allows elliptical symmetry about the mean. This distribution has been studied by Bingham (1964) and is introduced in the following section.

### Bingham's Distribution

Bingham's distribution has the form

$$\exp\left[\zeta_1(\mu_1'\mathbf{x})^2+\zeta_2(\mu_2'\mathbf{x})^2+\zeta_3(\mu_3'\mathbf{x})_2^2\right]ds/4\pi\,{}_1F_1(1/2;3/2;Z) \qquad (3)$$

where $\zeta_1$, $\zeta_2$, $\zeta_3$ are dispersion parameters and, using Kiraly's (1969) terminology for $\mu_3$, $\mu_1$, and $\mu_2$, $\mu_3$ is the best axis, $\mu_1$ is the axis of best zone, and $\mu_2$ is the axis of a zone containing $\mu_1$ and $\mu_3$. Also,

$$Z=\begin{bmatrix}\zeta_1 & 0 & 0\\ 0 & \zeta_2 & 0\\ 0 & 0 & \zeta_3\end{bmatrix}$$

and ${}_1F_1$ (1/2; 3/2; $Z$) is a hypergeometric function of matrix argument. Bingham imposes the constraint $\zeta_3 = 0$ to render the maximum likelihood estimate of $Z$ unique.

If the data are concentrated about a preferred orientation, both $\zeta_1$ and $\zeta_2$ will be negative. If $\zeta_1 \neq \zeta_2$, then $\zeta_1 < \zeta_2$ and the contours of the distribution will be elliptical in shape. In this example, the axis $\mu_1$ would be parallel to the "minor axis" of the "ellipse," and the axis $\mu_2$ would be parallel to the "major axis" of the "ellipse." The axis $\mu_3$, that is, the best axis or the mean of the distribution, would be perpendicular to both $\mu_1$ and $\mu_2$.

Estimation of parameters ($\zeta_i$, $\mu_i$, $i$ = 1,2,3) is discussed by Bingham but will be briefly touched on here. Given a set of orientation data

$$C=\{(\theta_1,\phi_1),(\theta_2,\phi_2),\ldots,(\theta_n,\phi_n)\}$$

define $\mathbf{x}_i$ by

$$\mathbf{x}_i=\begin{bmatrix}\sin\phi_i\cos\theta_i\\ \sin\phi_i\sin\theta_i\\ \cos\phi_i\end{bmatrix}$$

The coordinates of the vector $\mathbf{x}_i$ are the rectangular coordinates of a point on the unit sphere whose spherical coordinates are (1, $\theta_i$, $\phi_i$) (see Fig. 1).

Further, define a $3\times n$ matrix $X$ of column vectors by

$$X=[\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_n]$$

Let $\omega_1$, $\omega_2$, and $\omega_3$ be the eigenvectors of the matrix

$$XX^T=\begin{bmatrix}\Sigma x_i^2 & \Sigma x_iy_i & \Sigma x_iz_i\\ \Sigma x_iy_i & \Sigma y_i^2 & \Sigma y_iz_i\\ \Sigma x_iz_i & \Sigma y_iz_i & \Sigma z_i^2\end{bmatrix}$$

where $\omega_1<\omega_2<\omega_3$. Then the corresponding eigenvectors are the maximum likelihood estimates, $\hat{\mu}_1$, $\hat{\mu}_2$, and $\hat{\mu}_3$, of $\mu_1$, $\mu_2$, and $\mu_3$. Estimates of $\zeta_1$ and $\zeta_2$ are obtained by interpolating in tables given by Bingham (1964), who also has provided the information necessary for extending these tables.

According to Bingham (1964, p. 49), "$\sigma_{ij}{}^2$ is the variance of the rotation of $\hat{M}$ (considered as a set of three axes in space) about the axis $\mu_k$, where $i \neq k \neq j$." Here, $\hat{M} = [\hat{\mu}_1, \hat{\mu}_2, \hat{\mu}_3]$ and

$$\sigma_{ij}{}^2 = 1/2(\sigma_i - \sigma_j)(\omega_i - \omega_j)$$

The authors have developed a computer code which computes $\zeta_1$, $\zeta_2$, $\hat{\mu}_1$, $\hat{\mu}_2$, $\hat{\mu}_3$, $\sigma_{21}$, $\sigma_{31}$, and $\sigma_{32}$ and performs the $\chi^2$ test described in the following section.

## The Chi-Square Test

In this section $(\theta, \phi)$ will denote the spherical coordinates of a unit vector computed with respect to $\hat{\mu}_3$.

The number of points expected in the spherical quadrilateral

$$\{(\theta, \phi) | \cos\phi_2 \leqslant \cos\phi \leqslant \cos\phi_1, \theta_1 \leqslant \theta \leqslant \theta_2\}$$

is

$$[n/4\pi\, {}_1F_1(1/2; 3/2; Z)] \int_{\theta_1}^{\theta_2} \int_{\phi_2}^{\phi_1} \exp[(\zeta_1 \cos^2\theta + \zeta_2 \sin^2\theta)\sin^2\phi] \sin\phi d\phi d\theta \qquad (4)$$

According to Bingham (1964, p. 43), this is the most natural method of setting up class intervals if the data are concentrated strongly about $\hat{\mu}_3$, which is the situation in the sets of orientation data studied by the authors.

To apply the $\chi^2$ test to the set of orientation data $C$ (previous section), we first compute $\hat{\zeta}_1$, $\hat{\zeta}_2$, $\hat{\mu}_1$, $\hat{\mu}_2$, and $\hat{\mu}_3$. We then map the set $C$ onto the unit hemisphere by means of a linear transformation which maps the three axes $\hat{\mu}_1$, $\hat{\mu}_2$, and $\hat{\mu}_3$ onto the $x$, $y$, and $z$ coordinate axes, respectively. The image of $C$ is a set of points clustered about a point having (0,0) as its spherical coordinates and (0,0,1) as its rectangular coordinates. Next, we compute the spherical coordinates of the points in the image of $C$. Note that $\hat{\mu}_3$ is represented by (0,0,1) in the transformed system.

Class intervals are constructed in such a manner that each interval has an expected value greater than or equal to 5. Geometrically, the unit hemisphere, in the transformed coordinate system, is partitioned into a set of concentric circular bands, centered at $\hat{\mu}_3$, each of which is divided into four class intervals. The number of points in each class interval is easily counted. The expected value for each class interval is computed by using quadrature formulae to evaluate expression (4). The $\chi^2$ value thus obtained then is compared with tables of the $\chi^2$ distribution.

## NUMERICAL EXAMPLE

The technique of delineation and analysis of clusters in orientation data described previously was applied to a set of fracture orientations measured in

a porphyry copper deposit (San Manual mine, Arizona). The data are listed in Table 1, which is divided into three parts corresponding to the three clusters in the partition defined for $r = 0.0874$. Figure 2 shows polar equal-area projection of the data of Table 1; an illustrative demarcation of the boundaries of the three clusters is sketched by hand.

The objective function $F(P)$, as seen in eq (2), assumed the value 180.61 at the partition corresponding to $r = 0.0874$; this was the smallest value assumed over the range $0.0195 \leqslant r \leqslant 0.187$. A plot of $F(P)$ versus $r$ is shown in Figure 3.

As noted earlier, the objective function offered here was tested on generated data (prior to its application to the numerical example) and consistently minimized at partitions whose clusters were close to the generated clusters. In the example of fracture orientation data treated here, three distinct groupings of dense points were observed for a wide range of $r$ values: $0.0437 \leqslant r \leqslant 0.187$ (Fig. 4 shows polar equal-area projection of dense points corresponding to

**Table 1. Partition Elements for Numerical Example[a]**

| Cluster No. 1 | | | | | |
|---|---|---|---|---|---|
| θ,ϕ | θ,ϕ | θ,ϕ | θ,ϕ | θ,ϕ | θ,ϕ |
| 28,90 | 36,88 | 37,85 | 41,87 | 44,90 | 45,87 |
| 50,90 | 50,90 | 53,87 | 56,87 | 57,87 | 67,87 |
| 74,90 | 86,86 | 93,86 | 109,85 | 116,86 | 220,86 |
| 230,86 | 230,85 | 237,87 | 240,85 | 240,85 | 240,87 |
| 243,88 | 250,87 | 250,87 | 251,86 | 252,86 | 255,87 |
| 257,86 | 258,87 | 270,90 | 270,85 | 270,88 | 274,87 |
| 280,90 | 286,90 | 287,90 | 296,90 | 30,80 | 45,83 |
| 62,83 | 65,80 | 73,83 | 75,80 | 83,82 | 84,80 |
| 85,80 | 101,80 | 105,80 | 110,82 | 210,80 | 232,82 |
| 232,82 | 234,82 | 244,81 | 248,82 | 255,82 | 257,80 |
| 264,80 | 266,84 | 276,82 | 291,80 | 297,80 | 84,75 |
| 218,78 | 240,77 | 245,79 | 245,78 | 247,78 | 247,75 |
| 250,77 | 251,78 | 255,77 | 279,75 | 285,75 | 302,78 |
| 265,73 | 270,73 | 270,74 | 225,67 | 242,70 | 248,70 |
| 255,68 | 277,66 | 90,65 | 217,65 | 220,65 | 229,64 |
| 244,65 | 247,63 | 255,65 | 263,63 | 266,65 | 268,65 |
| 297,65 | 77,58 | 228,60 | 232,60 | 233,60 | 244,58 |
| 270,57 | 281,60 | 292,57 | 230,55 | 235,53 | 240,55 |
| 270,54 | 276,54 | 297,55 | 301,52 | 210,47 | 217,47 |
| 230,50 | 255,50 | 216,43 | 218,45 | 286,42 | 291,45 |
| 216,33 | 247,32 | 275,35 | 249,30 | 254,25 | 202,75 |
| 199,43 | | | | | |

**Table 1. Continued**

| Cluster No. 2 | | | | | |
|---|---|---|---|---|---|
| θ,ϕ | θ,ϕ | θ,ϕ | θ,ϕ | θ,ϕ | θ,ϕ |
| 7,90 | 10,90 | 12,86 | 13,90 | 13,85 | 14,87 |
| 15,90 | 24,87 | 136,87 | 141,86 | 143,87 | 147,85 |
| 147,85 | 150,87 | 162,88 | 163,85 | 167,86 | 175,87 |
| 182,86 | 314,90 | 314,88 | 315,90 | 325,86 | 326,90 |
| 327,90 | 328,90 | 341,90 | 342,88 | 342,85 | 343,85 |
| 345,85 | 345,87 | 345,88 | 345,87 | 346,86 | 350,86 |
| 350,90 | 353,85 | 355,90 | 356,87 | 356,85 | 358,87 |
| 360,88 | 13,82 | 150,83 | 155,80 | 186,83 | 186,83 |
| 327,80 | 327,83 | 328,80 | 330,84 | 331,80 | 322,82 |
| 336,82 | 338,80 | 342,80 | 350,83 | 352,80 | 355,80 |
| 355,84 | 192,77 | 343,60 | 328,78 | 346,79 | 351,75 |
| 354,75 | 354,75 | 355,75 | 357,78 | 360,77 | 334,72 |
| 344,74 | 356,74 | 357,74 | 6,70 | 8,70 | 10,68 |
| 15,67 | 140,67 | 147,70 | 161,70 | 304,70 | 314,70 |
| 333,67 | 334,66 | 337,70 | 337,70 | 341,70 | 344,70 |
| 344,70 | 346,70 | 353,70 | 360,70 | 360,70 | 10,64 |
| 13,63 | 193,65 | 340,65 | 347,62 | 350,65 | 9,56 |
| 28,58 | 149,60 | 167,57 | 313,58 | 328,60 | |
| 355,60 | 304,53 | 8,47 | 16,34 | 308,27 | |
| Cluster No. 3 | | | | | |
| θ,ϕ | θ,ϕ | θ,ϕ | θ,ϕ | θ,ϕ | θ,ϕ |
| 168,46 | 177,46 | 83,42 | 106,44 | 176,43 | 205,36 |
| 60,37 | 118,37 | 154,40 | 177,36 | 181,40 | 133,35 |
| 30,34 | 81,32 | 105,32 | 119,32 | 126,31 | 20,25 |
| 157,34 | 62,26 | 97,28 | 129,26 | 167,30 | 121,20 |
| 25,21 | 33,20 | 57,24 | 70,20 | 115,25 | 260,20 |
| 126,20 | 127,23 | 133,20 | 145,25 | 158,22 | 214,18 |
| 267,22 | 325,20 | 77,18 | 111,19 | 153,16 | 180,0 |
| 7,15 | 81,13 | 85,12 | 132,6 | 190,7 | |

[a] $(1, \theta, \phi)$ are the spherical coordinates of an observation (see Fig. 1).

the partition for $r = 0.0874$). Therefore, it was concluded that three predominant fracture sets were present in the porphyry copper deposit. However, the differences between any two of the ten partitions the authors examined (for the range of $r$ yielding three clusters) were small. For example, the maximum angular difference between a pair of best axes, $\hat{\boldsymbol{\mu}}_3$, for cluster 3 was 3.1°. The maximum angular differences for clusters 1 and 2 were 1.9° and 1.6°, respectively. Such differences, in general, will depend on the number of data

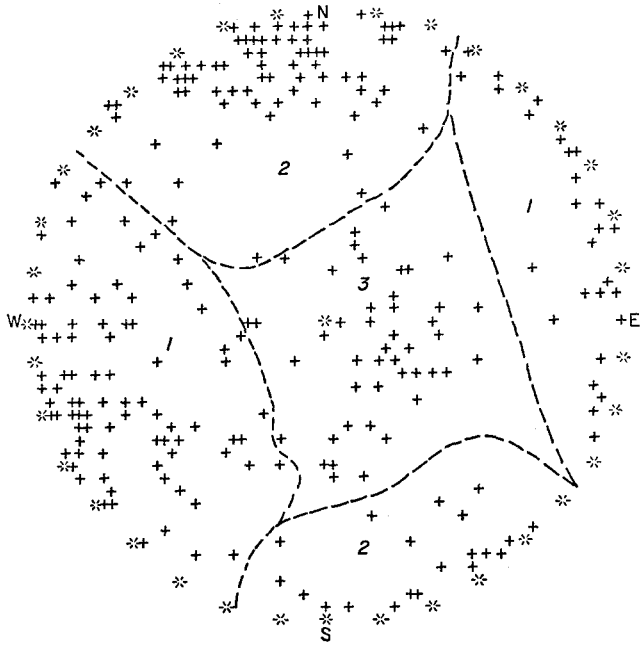



Figure 2. Polar equal-area projection of partitioned sample of Table 1. Digits 1, 2, and 3 denote clusters to which indicated data belong.

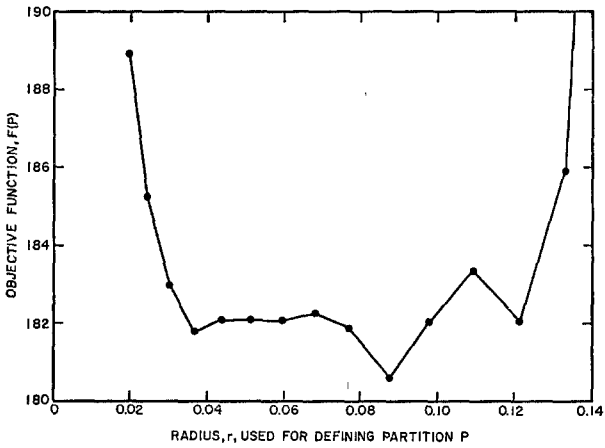



Figure 3. Plot of objective function versus $r$ for numerical example.

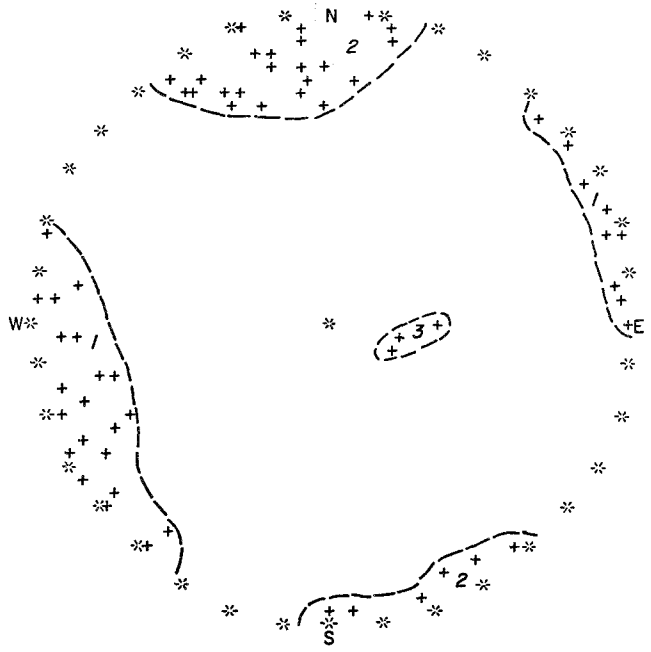



Figure 4. Polar equal-area projection of dense points for $r = 0.0874$; data from Table 1.

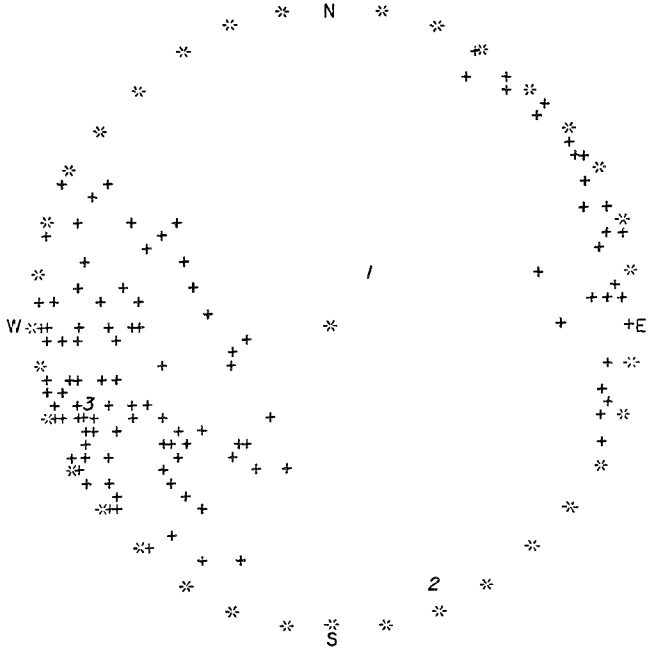



Figure 5. Polar equal-area projection of cluster 1. Digits 1, 2, and 3 denote vectors $\hat{\mu}_1$, $\hat{\mu}_0$, $\hat{\mu}_2$, respectively.

**Table 2. Maximum Likelihood Estimates of Parameters Associated with Bingham's Distribution for Data of Table 1**

| CLUSTER NUMBER | $\hat{\mu}_1$ | $\hat{\mu}_2$ | $\hat{\mu}_3$ | $\hat{\phi}_1$ | $\hat{\phi}_2$ | $\hat{\phi}_3$ | $\omega_1$ | $\omega_2$ | $\omega_3$ | $\sigma_{21}$ | $\sigma_{31}$ | $\sigma_{32}$ | $\hat{\tau}_1$ | $\hat{\tau}_2$ | 0.00 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 0.223 | -0.925 | 0.307 | 16.3 | 81.2 | 103.6 | 12.22 | | | 15.10 | 1.79 | 2.13 | -5.98 | | |
| 1 | 0.170 | 0.347 | 0.922 | 37.4 | 159.5 | 71.6 | | 17.07 | | | | | | -4.50 | |
| | 0.960 | 0.153 | -0.235 | | | | | | 97.71 | | | | | | 0.00 |
| | -0.217 | 0.183 | 0.959 | 19.4 | 73.0 | 80.9 | 7.10 | | | 17.84 | 1.47 | 1.70 | -8.67 | | |
| 2 | -0.251 | 0.939 | -0.235 | 229.2 | 79.0 | 346.2 | | 9.63 | | | | | | -6.64 | |
| | 0.943 | 0.292 | 0.157 | | | | | | 95.27 | | | | | | 0.00 |
| | 0.180 | 0.971 | 0.156 | 74.6 | 84.0 | 163.4 | 2.58 | | | 0.00 | 0.00 | 0.00 | -9.85 | | |
| 3 | -0.947 | 0.214 | -0.239 | 280.7 | 12.4 | 303.2 | | 4.64 | | | | | | -5.87 | |
| | 0.265 | 0.105 | -0.958 | | | | | | 39.78 | | | | | | 0.00 |

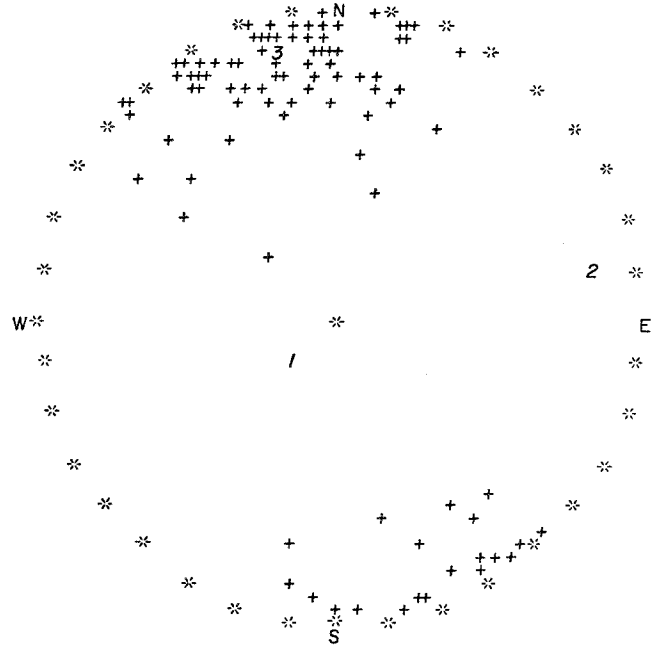



Figure 6. Polar equal-area projection of cluster 2. Digits 1, 2, and 3 denote vectors $\hat{\mu}_1$, $\hat{\mu}_2$, and $\hat{\mu}_3$, respectively.

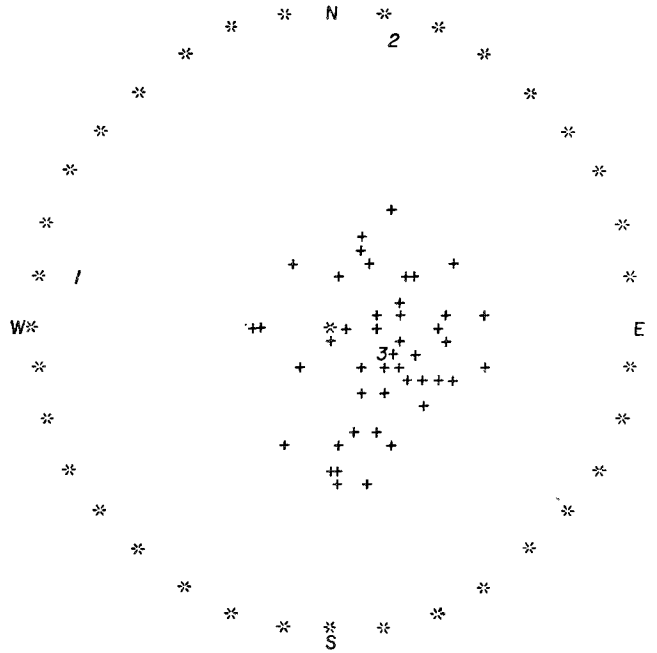



Figure 7. Polar equal-area projection of cluster 3. Digits 1, 2, and 3 denote vectors $\hat{\mu}_1$, $\hat{\mu}_2$, and $\hat{\mu}_3$, respectively.

**Table 3. Chi-Square Test for Three Clusters of Table 1**

| Cluster number | $n$ (degrees of freedom) | $\chi^2_{0.20}(n)$ | $\chi^2_{0.10}(n)$ | $\chi^2_{0.05}(n)$ | Actual $\chi^2$ value[a] |
|---|---|---|---|---|---|
| 1 | 18 | 22.760 | 25.989 | 28.869 | 27.73 |
| 2 | 14 | 18.151 | 21.064 | 23.685 | 17.12 |
| 3 | 2 | 3.219 | 4.605 | 5.991 | 4.90 |

[a] Expected value per class interval was between 5 and 6.5 for each cluster.

points in the given clusters as well as on the scatter of these points. It is noted that some care and good judgment, based on previously established characteristics of the orientation data that relate to the problem analyzed, should be exercised when using an objective function for selecting a desired or "best" partition.

For the numerical example (data of Table 1), equal-area projections of clusters 1, 2, and 3 corresponding to the partition defined for $r = 0.0874$ are given in Figures 5, 6, and 7, respectively. The maximum likelihood estimates of the parameters associated with Bingham's distribution, and computed from the data points belonging to the three clusters, are presented in Table 2. The results of the $\chi^2$ test for the three clusters are given in Table 3; it can be seen that the test is nearly significant, for all three clusters of this example, at the 95 percent level.

## SUMMARY

A technique is presented for clustering orientation data and for testing the resulting clusters against a probability distribution defined on the unit sphere which allows elliptical symmetry about its mean. Several partitions of a given sample of orientations are defined by the one-level mode analysis of Wishart. An objective function, $F(P)$, is evaluated at each partition, $P$. The clusters belonging to a partition near the point at which $F(P)$ assumes its minimum are compared with Bingham's probability distribution using the chi-square test. An application of the technique is made to an example data set composed of orientations of fractures in a porphyry copper deposit. Three clusters were defined in the example data set corresponding to the "best" partition selected by using $F(P)$. The results of analyses of the numerical example are listed and polar equal-area projections of the example data set as well as of the points in the three clusters are presented.

## ACKNOWLEDGMENTS

The authors wish to thank R. J. Alldredge of College of Agriculture, Uni-

versity of Nevada, Reno, Nevada, and D. M. Ross-Brown of Dames and Moore, Consulting Engineers, Los Angeles, California, for a constructive review of the manuscript of the paper.

## REFERENCES

Bingham, C., 1964, Distributions on the sphere and on the projective plane: unpubl. doctoral dissertation, Yale Univ., 93 p.

John, K. W., 1968, Graphic stability analysis of slopes in jointed rock: Jour. Soil Mech. and Found. Div., Proc. Am. Soc. Civil Engineers, v. 94, no. SM 2, p. 497–526.

Kiraly, L., 1969, Statistical analysis of fractures (orientation and density): Geologische Rundschau, v. 59, no. 1, p. 125–151.

Mahtab, M. A., 1974, Influence of natural jointing on coal mine stability and on the preferred direction of mine layout: BuMines Info. Circ. 8630, p. 70–78.

Mahtab, M. A., and Dixon, J. D., 1975, Influence of rock fractures and block boundary weakening on cavability: AIME Annual Meeting, New York, preprint no. 75-AM-9.

Mahtab, M. A., Bolstad, D. D., Alldredge, J. R., and Shanley, R. J., 1972, Analysis of fracture orientations for input to structural models of discontinuous rock: BuMines Rept. Invest. 7669, 76 p.

Shanley, R. J., and Mahtab, M. A., 1975, FRACTAN: A computer code for analysis of clusters defined on the unit hemisphere: BuMines Info. Circ. 8671, 49 p.

Wishart, D., 1968, Mode analysis: A generalization of nearest neighbour which reduces chaining effects, *in* Numerical taxonomy: Academic Press, London and New York, p. 282–309.