A UNITED STATES
DEPARTMENT OF
COMMERCE
PUBLICATION

NOAA Technical Report NOS 44

U.S. DEPARTMENT OF COMMERCE
National Oceanic and Atmospheric Administration
National Ocean Survey

The Determination of Focal Mechanisms Using P- and S-Wave Data

WILLIAM H. DILLINGER
ALLEN J. POPE
SAMUEL T. HARDING

ROCKVILLE, MD.

July 1971

NOAA TECHNICAL REPORTS

National Ocean Survey Series

The National Ocean Survey (NOS) provides charts and related information for the safe navigation of marine and air commerce. The Survey also furnishes other earth science data--from geodetic, hydrographic, oceanographic, geomagnetic, seismologic, gravimetric, and astronomic surveys or observations, investigations, and measurements--to protect life and property and to meet the needs of engineering, scientific, defense, commercial, and industrial interests.

Because many of these reports deal with new practices and techniques, the views expressed are those of the authors, and do not necessarily represent final Survey policy. NOS series of NOAA Technical Reports is a continuation of, and retains the consecutive numbering sequence of, the former series, ESSA Technical Report Coast and Geodetic Survey (C&GS), and the earlier series, C&GS Technical Bulletin.

Reports in the series are available through the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 20402. Those publications marked by an asterisk are out of print.

Coast and Geodetic Survey Technical Bulletins

*C&GS 1. Aerotriangulation Adjustment of Instrument Data by Computational Methods. William D. Harris, January 1958 (Superseded by No. 23.)
*C&GS 2. Tellurometer Traverse Surveys. Lt. Hal P. Demuth, March 1958.
*C&GS 3. Recent Increases in Coastal Water Temperature and Sea Level--California to Alaska. H. B. Stewart, Jr., B. D. Zetler, and C. B. Taylor, May 1958.
*C&GS 4. Radio Telemetry Applied to Survey Problems. Richard R. Ross, February 1959.
*C&GS 5. Raydist on Georges Bank. Capt. Gilbert R. Fish, April 1959.
C&GS 6. The Tsunami of March 9, 1957, as Recorded at Tide Stations. Garrett G. Salsman, July 1959. Price $0.15.
*C&GS 7. Pantograph Adjustment. G. C. Tewinkel, July 1959.
*C&GS 8. Mathematical Basis of Analytic Aerotriangulation. G. C. Tewinkel, August 1959 (Superseded by No. 21.)
*C&GS 9. Gravity Measurement Operations in the Field. Lt. Comdr. Hal P. Demuth, September 1959.
*C&GS 10. Vertical Adjustment of Instrument Aerotriangulation by Computational Methods. William B. Harris, September 1959 (Superseded by No. 23.)
*C&GS 11. Use of Near-Earth Satellite Orbits for Geodetic Information. Paul D. Thomas, January 1960.
*C&GS 12. Use of Artificial Satellites for Navigation and Oceanographic Surveys. Paul D. Thomas, July 1960.
C&GS 13. A Singular Geodetic Survey. Lansing G. Simmons, September 1960. Price $0.15.
*C&GS 14. Film Distortion Compensation for Photogrammetric Use. G. C. Tewinkel, September 1960.
*C&GS 15. Transformation of Rectangular Space Coordinates. Erwin Schmid, December 1960.
*C&GS 16. Erosion and Sedimentation--Eastern Chesapeake Bay at the Choptank River. G. F. Jordan, January 1961.
C&GS 17. On the Time Interval Between Two Consecutive Earthquakes. Tokuji Utsu, February 1962. Price $0.10.

(Continued on inside back cover)

U.S. DEPARTMENT OF COMMERCE
Maurice H. Stans, Secretary

NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION
Robert M. White, Administrator

NATIONAL OCEAN SURVEY
Don A. Jones, Director

NOAA Technical Report NOS 44

The Determination of Focal Mechanisms Using P- and S- Wave Data

WILLIAM H. DILLINGER
ALLEN J. POPE
SAMUEL T. HARDING

OFFICE OF SEISMOLOGY AND GEOMAGNETISM
OFFICE OF GEODESY AND PHOTOGRAMMETRY

ROCKVILLE, MD

JULY 1971

UDC 550.341.063:550.344.094:550.340.15:681.3

550.34	Seismology
.341.063	Methods of determining the location of origin of earthquakes
.344.094	Behavior of waves, properties, and processes
.340.15	Graphical representation of computation methods
681.3	Data processing with computers

For sale by the Superintendent of Documents, Government Printing Office
Washington, D.C., 20402 - Price 60 cents
Stock Number 0321-0005

CONTENTS

Page

Abstract..... 1

I. Introduction..... 2

II. S-wave, P-wave, and P- and S-wave mechanism solutions 3

A. Theory of the solutions..... 3

B. S-wave solution..... 3

C. P-wave solution..... 7

D. Combined P- and S-wave solution..... 11

E. Focal plane determination..... 14

III. Exact P-wave limits..... 20

IV. S-wave and P-wave fiducial regions..... 25

A. Theory of the fiducial regions..... 25

B. S-wave fiducial regions..... 32

C. P-wave fiducial regions..... 34

V. Data processing..... 42

VI. Conclusion..... 50

Acknowledgments..... 52

References..... 52

List of notations..... 55

LIST OF FIGURES

Page

Figure 1.--Perspective view and corresponding lower hemisphere representation of the focal planes..... 16

Figure 2.--Relation between the station complementary circles and the poles of the planes which define the compressional and dilatational quadrants..... 21

Figure 3.--Construction of the exact boundaries for the maximum-likelihood P-wave solution..... 24

Figure 4.--A comparison between the computer-contoured and the exact boundaries for the maximum-likelihood P-wave solution..... 26

Figure 5.--Computer plot showing P-wave first motion data on the lower hemisphere..... 35

Figure 6.--Computer plot showing observed S-polarization angles on the lower hemisphere..... 37

Figure 7.--Computer plot of the predicted S-polarization angles for the combined P- and S-wave solution..... 39

Figure 8.--System flow chart showing the operations and computer programs involved in computing focal mechanisms.. 41

Figure 9.--Computer plot showing the 96.5-percent fiducial limit for the P-wave solution..... 45

Figure 10.--Computer plot of the 95-percent fiducial limits for the S-wave solution..... 47

Figure 11.--Computer plot showing the best combined P- and S-wave solution and contours of equal likelihood around the maximum-likelihood solution..... 49

Figure 12.--Computer plot showing a profile of the S-wave score with the initial plane held fixed at the value given..... 51

THE DETERMINATION OF FOCAL MECHANISMS USING P- AND S-WAVE DATA

William H. Dillinger[1], Allen J. Pope[2], and Samuel T. Harding[1]

ABSTRACT. The National Oceanic and Atmospheric Administration computes focal plane solutions using P-wave first motion, S-wave polarization angles, and combinations of the two. These calculations have been programmed for a CDC-6600 computer. A new method is used, based on a maximum-likelihood argument implemented by the generation of a rigorously derived score surface on a discrete grid. The combined score is of the form a(P-score) + b(S-score) where a and b and the score values are derived by a maximum-likelihood argument. For the P-solution and S-solution, the computer system also produces contoured fiducial regions about the poles of the focal planes on an equal-area projection of the lower focal hemisphere. This permits an evaluation of the quality of the solutions and makes possible the comparison of solutions for different earthquakes. These contoured limits have a probabilistic interpretation as 95-percent fiducial regions, a term explained in the context of the general problem of the construction of confidence regions. A new graphical method for P-wave solutions makes possible the construction of exact boundaries of the regions of possible variation. The Aleutian earthquake of May 14, 1969, is analyzed, and the solution shows good agreement between the observed P-wave first motions and the S-polarization angles. This analysis gives practical support to the validity of combined solutions.

---

[1]Geophysics Research Group, Office of Seismology and Geomagnetism, NOS.

[2]Geodetic Research and Development Laboratory, Office of Geodesy and Photogrammetry, NOS.

I. INTRODUCTION

A number of computerized techniques for using P-waves, S-waves, or both, to determine focal mechanisms have been developed in recent years. Programs have been developed for P-wave solutions by Knopoff (1961), Kasahara (1963), Wickens and Hodgson (1967), and Zhelankina et al. (in press). Programs for the S-wave solutions have been developed by Udias (1964), Hirasawa (1966), and Stevens (1967) using S-wave polarization angles, and by Zhelankina et al. using the direction of SH and SV first motions. These solutions have dealt with the P-wave and S-wave data separately. A more recent proposal for a combined P- and S-solution (Udias and Baumann 1969) lacks the statistical support that permits incorporation of objective evaluations of the quality of P- and S-data.

Increased use of computers in the determination of focal mechanisms has helped to point out large discrepancies between solutions given by different observers and has indicated that a number of these solutions must be regarded as marginal at the very best. This has been documented by Wickens and Hodgson (1967) and by Stevens and Hodgson (1968).

Sykes (1967), Stauder (June 1968, December 1968) and Isacks et al. (1968) have demonstrated the value of mechanism solutions in the study of regional and global tectonic patterns. A method permitting the properly qualified use of marginal solutions and giving an objective and easily interpreted indication of the quality of solutions will undoubtedly be useful in studying these patterns. The purpose of this paper is to present a method for finding and visually and statistically evaluating mechanism solutions.

The method consists of an interpretation of the error surfaces generated by plotting as a function of the model parameters (e.g., strike and dip of the nodal planes) an appropriate measure of the discrepancy between the observed P-wave first motions, S-wave polarization angles, or both, and those predicted on the basis of a particular focal mechanism model. These surfaces show that it is usually difficult to arrive at a unique set of parameter values because there is typically a range of acceptable solutions. A good solution is characterized by a narrow range of acceptable parameter values, that is, by small fiducial regions; a poor solution is characterized by large, ill-formed fiducial regions.

II. S-WAVE, P-WAVE, AND P- and S-WAVE MECHANISM SOLUTIONS

A. Theory of the Solutions

Source mechanism determinations involve nonlinear mathematical models and, in the case of P-wave first motion data, nonnormal statistics. To pursue the implications of these features, we must first define what is meant by a solution.

The best focal mechanism solution for given data and a given model of the source mechanism may be defined as that optimizing some agreed-upon measure, such as the sum of squares, of the difference between observed quantities and their value predicted from the model. However, because a sound choice of the measure of agreement and the construction of correctly scaled fiducial regions about the resulting solution are rooted in the statistics of the observations, it is desirable to return to a basic criterion which uses these statistics.

The method of maximum likelihood produces estimates having many desirable properties. The statistics of the observations are fully used, in contrast to the minimum variance method, for example, which uses only the second moments of the distribution of the observational errors. The problem at hand involves both continuous data (S-wave polarization angles), discrete data (P-wave first motions), and combinations of the two. The statistical argument leading to maximum-likelihood solutions and fiducial regions differs in these three cases. Details of the numerical procedure for the construction of P-wave fiducial regions are sensitive to subtle variations in the statistical argument. In addition, various standard simplifying assumptions that play an important role in most applications of statistics to adjustment problems no longer apply. Thus, the statistical discussion leading to the solutions and fiducial regions is expanded in an attempt to convey some appreciation of this aspect of the problem. We first describe the formulation of the maximum-likelihood solutions and the numerical procedure for finding these solutions before discussing the fiducial regions.

B. S-Wave Solution

Seismologically, P-wave first motion data can be said to be more important than S-wave data because of their greater density, ease of measurement, and reliability. However, for an exposition of the statistics involved, S-wave data are logically prior to P-wave data because S-wave data depart least from the ordinary adjustment situation in which statistical considerations support a least-squares criteria. We will recapitulate this development to provide a framework for the derivation of the P-wave first motion maximum-likelihood solution.

Suppose we are given $\overline{\chi}$, a vector of observable random variables defined on a continuous sample space, and the form of the associated joint probability density function (p.d.f.) conditional upon a vector of unknown parameters $\overline{\beta}$. This p.d.f. is denoted by $\phi(\overline{\chi}|\overline{\beta})$. The vector of observables $\overline{\chi}$ has as its ith component the observed polarization angle at station i, ε_i. The vector of parameters $\overline{\beta}$ is formed from the variables necessary to specify that part of the model of the earthquake mechanism that affects $\overline{\chi}$. The expected values of the observables are expressed as nonlinear functions of $\overline{\beta}$ by the assumed model of the earthquake mechanism. For the purpose of this analysis, the difference between observed values and those predicted from the postulated model is attributed solely to the presence of random errors whose p.d.f. is given by ϕ.

The associated p.d.f. $\phi(\overline{\chi}|\overline{\beta})$ can be thought of either as the probability density at the n-dimensional point $\overline{\chi}$ or as the joint probability density of the n observables χ_i, $\phi(\chi_1, \chi_2, \ldots, \chi_n|\overline{\beta})$. The vertical bar in $\phi(\overline{\chi}|\overline{\beta})$ means "conditional upon." The notation $\phi(\overline{\chi},\overline{\beta})$, on the other hand, denotes the joint p.d.f. of the random variables $\overline{\chi}$ and $\overline{\beta}$. By definition, $\phi(\overline{\chi}|\overline{\beta}) = \phi(\overline{\chi},\overline{\beta})/\phi(\overline{\beta})$ where $\phi(\overline{\beta})$ is the marginal p.d.f. of $\overline{\beta}$, $\phi(\overline{\beta}) = \int_{\text{all } \chi}\phi(\overline{\chi},\overline{\beta})d\overline{\chi}$. Thus, the use of this notation suggests that $\overline{\beta}$ in $\phi(\overline{\chi}|\overline{\beta})$ consists of particular fixed values of a random variable having an associated density $\phi(\overline{\beta})$. This corresponds to a Bayesian approach in which parameters are interpreted as random variables both before and after the introduction of data. The "a priori" p.d.f. of β, $\phi(\overline{\beta})$ often expresses a subjective "degree of ignorance" about $\overline{\beta}$; the "a posteriori" p.d.f. of $\overline{\beta}$, $\phi(\overline{\beta}|\overline{\chi})$ is obtained from Bayes' formula, $\phi(\overline{\beta}|\overline{\chi}) = \phi(\overline{\chi}|\overline{\beta})\phi(\overline{\beta})/\phi(\overline{\chi})$. An alternate notation $\phi(\overline{\chi};\overline{\beta})$, read as "the p.d.f. of $\overline{\chi}$ functionally dependent upon fixed but initially unknown parameters $\overline{\beta}$," is appropriate to a strict non-Bayesian point of view in which subjective prior distributions (as opposed to distributions supported by empirical evidence) are excluded. $\overline{\beta}$ is regarded as fixed but initially unknown, or equivalently, the prior density of $\overline{\beta}$ must be a Dirac delta with unknown point of support. This distinction can be ignored for solutions based on maximum-likelihood estimation defined below. However, it is relevant in the interpretation of fiducial regions for these solutions. As for the long history of debate surrounding this point, suffice it to say that in geophysics and lately in engineering practice as a whole (Benjamin and Cornell 1970), the Bayesian point of view is favored because of the need for statistical judgments in situations where a non-Bayesian approach is operationally unrealizable either because of the shortcomings of the data (e.g., "one-shot" situations where the strict frequency interpretation of probability becomes empirically untestable) or because of the practical

difficulty or impossibility of non-Bayesian constructions (e.g., the "strict confidence regions" described below).

When considered as a function of variable $\overline{\beta}$ given particular observed values of $\overline{\chi}$, $\phi(\overline{\chi}|\overline{\beta})$ is called the likelihood of $\overline{\beta}$ and is denoted by $L(\overline{\beta};\overline{\chi})$. Then, the maximum-likelihood estimate of $\overline{\beta}$ based on the particular observed $\overline{\chi}$ is defined as the value of $\overline{\beta}$ maximizing $L(\overline{\beta};\overline{\chi})$ and is denoted by $\hat{\beta}_{ML}$. From the modern statistical viewpoint, this definition is justified by the many desirable properties of the estimates to which it gives rise rather than by a derivation from more basic principles (Rao 1965). This point of view avoids the difficulties attendant upon older attempts to prove maximum likelihood using Bayesian arguments. However, these difficulties are still present in the interpretation of fiducial regions.

The use of a different notation for likelihood and p.d.f. emphasizes the different domain of the two functions. The function $L(\overline{\beta};\overline{\chi})$ is defined over all $\overline{\beta}$ for a given $\overline{\chi}$, whereas $\phi(\overline{\chi}|\overline{\beta})$ is defined over all $\overline{\chi}$ for a fixed $\overline{\beta}$. This notation also emphasizes the fact that likelihood is not generally a p.d.f. over $\overline{\beta}$, although we shall see below in connection with fiducial regions that it can be renormalized and interpreted as a kind of p.d.f. over $\overline{\beta}$ under certain assumptions.

For the S-wave polarization angle data $\overline{\chi}_s$, the errors (i.e., the differences between the true and observed values of polarization angles) are assumed to be jointly normally distributed so that we have

$$L(\overline{\beta};\overline{\chi}_s) = \phi(\overline{\chi}_s|\overline{\beta}) = c \exp\{-\frac{1}{2\sigma^2} R_s\}$$

where c is a constant, $c = (\det W)^{1/2}(2\pi\sigma^2)^{-n/2}$, and $R_s = \overline{V}^t W \overline{V}$. Here $\overline{V}$ is the vector of S-wave polarization angle residuals $v_i = \varepsilon_i(\overline{\beta}) - \varepsilon_i^o$; that is, the ith residual is the difference between the expected value of the polarization angle $\varepsilon_i(\overline{\beta})$ at station i computed on the basis of model parameters $\overline{\beta}$ and the observed polarization angle at station i, ε_i^o. The $W = M^{-1}\sigma^2$ is a weight matrix where M is the covariance matrix of the observations and σ is the standard error of one polarization angle determination. In this work, polarization angle determinations at different stations are assumed to be uncorrelated and of equal standard error so that W = I, the unit matrix. Only slight modifications are needed to handle unequal standard errors. Thus, the quadratic form $R_s = \overline{V}^t W \overline{V}$ becomes $R_s = \sum_i v_i^2$, and

$$L(\overline{\beta};\overline{\chi}_s) = c \exp\{-\frac{1}{2\sigma^2} \sum_i v_i^2\}.$$

$L(\overline{\beta};\overline{\chi}_s)$ is maximized as a function of $\overline{\beta}$ for that $\overline{\beta}$ producing the the smallest $R_s = \sum_i v_i^2$, illustrating the well-known result that maximum-likelihood estimation applied to normally distributed data leads to a least-squares solution.

The parameter σ^2 appearing in L, the "variance of an observation of unit weight," can be estimated from the vector of least-squares residuals $\overline{V}_{LS}$ by the formula from linear least squares (see, e.g., Hamilton 1964, Rao 1965, and Draper and Smith 1966), so that

$$\hat{\sigma}^2 = \overline{V}_{LS}^t \, W\overline{V}_{LS}/(n - u)$$

or

$$\hat{\sigma}^2 = R_s^{LS}/(n - u)$$

where n is the number of observations and u is the number of parameters. Use of this estimate in the nonlinear case is justified below in the derivation of S-wave fiducial regions.

Note that $\overline{V}$ in $\overline{V}^t W\overline{V}$ and $\hat{\sigma}^2$ consists of residuals on polarization angles rather than functions of polarization angles and parameters as in Udias (1964). Alternative assignments of residuals give simpler analytical solutions. However, the direct numerical maximum-likelihood solution used here permits residuals to be put where required by the statistical analysis.

In addition, the direct numerical solution avoids difficulties in the solution arising from the nonlinearity of the model such as those caused by the presence of secondary minima in $R_s(\overline{\beta})$. The complete likelihood surface is explored and exhibited in a consistent manner for P- , S- , and P- and S-waves. Any stationary points not corresponding to solutions pose no problem as they may in Newton-Gauss (linearization--least-squares solution--iteration) solutions.

The polarization angles , ε, are the final result of several component measurements and judgments, each subject to error. The central limit theorem asserts that the distribution of a linear combination of n errors from rather arbitrarily distributed populations tends toward normality as n increases. The precise distribution of ε would take into account the 2π-periodicity of ε and the π-ambiguity introduced because of the inability to pick the sense of first motion in about 20 percent of polarization angle determinations. These details are neglected here in favor of a simplifying assumption of normality. Empirical studies of polarization angle residuals support the reasonableness of this approximation (Stauder and Bollinger 1964). A minimal interpretation is that maximum likelihood, based on

normality, is equivalent to the use of a least-squares criteria for the fit between model and observations. The specific form of the distribution is needed to obtain the fiducial regions which give a visual representation of the quality of a solution.

C. P-Wave Solution

The P-wave data consist of compressional or dilatational first motions at each station which either agree or disagree with the first motion predicted for that station by a model of the earthquake mechanism. Let P_i denote the probability that the observed first motion agrees with the true first motion at station i; that is, P_i is the probability of reading (and reporting!) the record correctly. Then the probability of disagreement, that is, of reading the record incorrectly, is $1 - P_i$. If there is evidence for significant variation in the quality of P-wave first motion determinations at various stations, this variation is correctly and entirely accounted for by variation in P_i. This variation in P_i amounts to a form of "weighting" appropriate for the P-wave observations. In fact, we use $P_i = P$ for all i. This is analogous to the use of the same standard errors at all stations (weight matrix = I) in the S-wave solution. In a more sophisticated solution, the quality of the observation as expressed in P_i could be related to the observed or predicted amplitude of the signal, but the assumption that $P_i = P$ has been found adequate. The equations below easily extend to the case of different P_i.

Some investigators have proposed "weighting" as a function of the separation of stations and nodal planes. In this report, weights are regarded as related only to the quality of the observations as expressed in P_i and M. The interaction of stations and nodal planes (e.g., the increased importance of stations near the nodal planes) is already present in a correctly formulated solution. Attempts to "guess in advance" or to arbitrarily influence the effect of specific data points are not logically well founded (also see comments below on the estimation of P).

An agreement or disagreement at station i is assumed statistically independent of agreement or disagreement at any other station j. Therefore, the probability of agreement at both stations i and j is P_iP_j, or the probability of agreement at station i and disagreement at station j is $P_i(1 - P_j)$. (The assumption of statistical independence was incorporated into the S-wave p.d.f. by the use of a diagonal covariance matrix M.)

Let χ_i^o denote the observed first motion at station i; that is, $\chi_i^o = +1$ for compression (away from the focus) and $\chi_i^o = -1$ for dilatation (toward the focus). Similarly, let $\chi_i^c(\overline{\beta}) = \pm 1$ denote the first motion predicted at station i by a model of the focal mechanism dependent on parameters $\overline{\beta}$. Then when the true sense of first motion at all stations is given by $\chi_i^c(\overline{\beta})$, i=1...N, the probability of observing a particular set of first motions χ_i^o, i = 1 ... N, can be written as

$$\prod_{i=1}^{N} \{\frac{1}{2} + (P_i - \frac{1}{2})\, \chi_i^o\, \chi_i^c\, (\overline{\beta})\} .$$

When χ_i^o and $\chi_i^c(\overline{\beta})$ agree, $\chi_i^o \chi_i^c = +1$ and the corresponding term of the product is P_i. If χ_i^o and χ_i^c disagree, $\chi_i^o \chi_i^c = -1$ and the corresponding term becomes $1 - P_i$. Thus in the product, there are R_p terms of the form P_i and $N - R_p$ terms of the form $1 - P_i$ where R_p is the number of agreements and $N - R_p$ is the number of disagreements (N is the total number of stations).

Let $\overline{\chi}_p$ denote the vector of observations having N-components χ_i^o. With $P_i = P$, the joint probability function of the observations $\overline{\chi}_p$ conditional upon the parameters $\overline{\beta}$, denoted by $Pr(\overline{\chi}_p|\overline{\beta})$, is the point binomial probability function

$$Pr(\overline{\chi}_p|\overline{\beta}) = P^{R_p} (1 - P)^{N - R_p} .$$

$Pr(\overline{\chi}_p|\overline{\beta})$ is a function of $\overline{\chi}_p$ and $\overline{\beta}$ because $R_p = R_p(\overline{\chi}_p, \overline{\beta})$. This function can be written explicitly in the form

$$R_p = \sum_{i=1}^{N} \{\frac{1}{2} + \frac{1}{2}\, \chi_i^o \chi_i^c(\overline{\beta})\} .$$

The vector of observations $\overline{\chi}_p$ consists of the one particular combination of compressions and dilatations observed at the N-stations from among the 2^N such combinations that are possible. Stated in another way, $Pr(\overline{\chi}_p|\overline{\beta})$ is the probability associated, for given values of the parameters $\overline{\beta}$, with the N-dimensional point $\overline{\chi}_p$ in the discrete sample space consisting of all 2^N possible combinations of compressions and dilatations that could be observed at N stations. A standard result from prob-

ability theory is that $\sum_{\overline{\chi}_p} \Pr(\overline{\chi}_p|\overline{\beta}) = 1$ where the sum is over all 2^N combinations.

For given $\overline{\chi}_p$, R_p exhibits "step-function" jumps when the parameters $\overline{\beta}$ change enough to cause a nodal plane to cross a station. It is this discontinuity of R_p as a function of $\overline{\beta}$ that makes difficult or impossible the usual solution-and-error analysis procedures that require evaluation of derivatives with respect to $\overline{\beta}$, such as Newton-Gauss (Taylor series) iteration, gradient methods, and large-sample M's of maximum-likelihood estimates. This situation makes the numerical method used here corrrespondingly more attractive, both in finding solutions and in constructing fiducial regions.

The probability function of $\Pr(\overline{\chi}_p|\overline{\beta})$ gives the probability of observing values $\overline{\chi}_p$ of the discrete observables when the parameters have the value $\overline{\beta}$, just as the probability element $\phi(\overline{\chi}_s|\overline{\beta})d\overline{\chi}_s$, formed from the p.d.f. ϕ, gives the probability of observing values of the continuous observables $\overline{\chi}_s$ in the range $d\overline{\chi}_s$ about $\overline{\chi}_s$. $\Pr(\overline{\chi}_p|\overline{\beta})$, thought of as a function of the parameters $\overline{\beta}$ for given observations $\overline{\chi}_p$, is defined as the likelihood of $\overline{\beta}$ and is denoted by $L(\overline{\beta};\overline{\chi}_p)$.

The maximum-likelihood estimate of $\overline{\beta}$ based on data $\overline{\chi}_p$, denoted by $\hat{\beta}_{ML}$, is that value of $\overline{\beta}$ maximizing $L(\overline{\beta};\overline{\chi}_p)$. There may, in fact, be an infinite set of $\overline{\beta}$'s, A (i.e., A is a region in the parameter space), such that $L(\overline{\beta}\epsilon A;\overline{\chi}_p) > L(\overline{\beta}\not\epsilon A;\overline{\chi}_p)$. In this case, A is referred to as the maximum-likelihood set estimator of $\overline{\beta}$. Note that L is a nondecreasing function of R_p for $P \geq .5$, and, therefore, is maximum when R_p is maximum. Thus, $\hat{\beta}_{ML}$ is the value of $\overline{\beta}$ producing the largest number of agreements with the observed first motions.

The probability P, a parameter of the probability function Pr, can be estimated on the basis of the maximum-likelihood solution $\hat{\beta}_{ML}$. Of the 2^N mutually exclusive combinations comprising the sample space of $\overline{\chi}_p$, $\binom{N}{R}$ produce exactly R agreements where $\binom{N}{R} = N!/R!(N-R)!$, the number of combinations of N-things taken R at a time. (The subscript p on R has been dropped for convenience.) The probability of a particular combination of R agreements and N-R disagreements, from above, is $P^R(1-P)^{N-R}$. Therefore,

the probability of a score of R, conditional on $\overline{\beta}$ and the configuration of the N-stations, is given by the binomial probability function

$$Pr(R \text{ agreements}) = \binom{N}{R} P^R (1 - P)^{N-R} .$$

That is, for fixed $\overline{\beta}$ and stations, R is a binomially distributed random variable. As shown in, for example, Wilks (1962), Mood and Graybill (1963), and Rao (1965), the maximum-likelihood estimate of the parameter P in the binomial probability function, based on a sample of size N producing R agreements, is

$$\hat{P} = R/N.$$

Since for $\overline{\beta} = \hat{\beta}_{ML}$, we have $R = R_{max}$, an estimate of P <u>based on</u> $\hat{\beta}_{ML}$ is

$$\hat{P} = \frac{R_{max}}{N} ,$$

corresponding to the formula for $\hat{\sigma}^2$ based on R_s. $\hat{P}$ is discussed at more length than $\hat{\sigma}^2$, in part to emphasize the distinction between point binomial and binomial distributions.

If N is small (<20 or so) or the station distribution is poor, $\hat{P}$ computed as above may be unrealistically near one, resulting in regions that are too small. In this case, a more realistic value of P based on past experience, usually $.90 \leq P \leq .95$, is used rather than R_{max}/N.

If $R_1 < .5N$, it is possible to change $\overline{\beta}$ so that all predicted first motions reverse, thus changing the number of agreements to $R_2 = N - R_1 > .5N > R_1$. Thus with R' defined by

$$R' = \begin{cases} R & R \geq .5N \\ N-R & R \leq .5N \end{cases} ,$$

the exact probability function of R' is given by

$$Pr(R') = \begin{cases} 0 & R < .5N \\ Pr(R) & R = .5N. \\ Pr(R) + Pr(N - R) & R > .5N \end{cases}$$

Note $\sum_{\text{all } R'} Pr(R') = 1$. However, the P's and N's encountered in practice, $P \geq .75$ (based on $\hat{P}$) and $N \geq 5$, give a very small

probability for R<.5N. Therefore, Pr(R') differs insignficantly from Pr(R), and the second term is neglected.

It has been suggested that P-wave motions inconsistent with those at nearby stations be discarded (Zhelankina et al., in press). While this may produce the illusion of a better solution, in fact, discarding inconsistent P-wave motions adulterates the main source of information about P--the probability of reading a first motion correctly. $\hat{P}$ influences the size of the P-wave fiducial regions and should be estimated from the total available data. Rejections, if made at all, should be made after an initial solution using all data to estimate P. "Inconsistent" stations may also affect the shape of the fiducial regions. In fact, because an impartial assessment of quality (requiring both the size and shape of fiducial regions) is always a goal in this report, no rejections based on inconsistencies alone are made. The method of maximum likelihood and the presentation of fiducial contours based on the likelihood or score surface use all of the information in the data, good and bad, for the solution and for the fiducial regions.

Henceforth, $\hat{P}$ and $\hat{\sigma}^2$ are to be understood for P and σ^2 where needed. In the construction of S-wave fiducial regions, it is possible to allow for the additional source of random variation introduced by use of $\hat{\sigma}^2$ in place of σ^2 by standard methods carried over from linear least-squares theory. A corresponding development for the P-wave case is not attempted. It is doubtful that this "second-order" statistical consideration is important, but this is a direction in which the analysis could be extended.

D. Combined P- and S-Wave Solution

The maximum-likelihood formulation of a combined P- and S-wave solution uses N_p observed first motions $\overline{\chi}_p$ defined over a discrete sample space and N_s polarization angles $\overline{\chi}_s$ defined on a continuous sample space.

In review, the likelihood function for S-wave data alone was

$$L(\overline{\beta};\overline{\chi}_s) = \phi(\overline{\chi}_s|\overline{\beta}) = c \exp\{-\frac{1}{2\sigma^2} R_s\}$$

and the likelihood function for discrete P-wave data was

$$L(\overline{\beta};\overline{\chi}_p) = Pr(\overline{\chi}_p|\overline{\beta}) = P^{R_p} (1 - P)^{N-R_p}$$

where $R_s = \overline{V}^t W \overline{V}$ and R_p = the number of agreements with the motion predicted by parameters $\overline{\beta}$. For convenience, R_s and R_p are denoted as the "S-score" and "P-score," respectively. The P-wave solution required R_p to be maximum and the S-wave solution required

R_s to be minimum. For the purpose of the combined solution, it is convenient to use $\overline{R}_p = N - R_p$, the number of disagreements, rather than R_p so that both $\overline{R}_p$ and R_s decrease for better solutions. Then

$$L(\overline{\beta};\overline{\chi}_p) = P^{N-\overline{R}_p} (1 - P)^{\overline{R}_p} .$$

Assuming statistical independence of errors in P-wave and S-wave data,

$$\Pr\{(\overline{\chi}_s \varepsilon d\overline{\chi}_s) \text{ and } (\overline{\chi}_p)|\overline{\beta}\} = \Pr(\overline{\chi}_p|\overline{\beta})\ \phi(\overline{\chi}_s|\overline{\beta})\ d\chi_s$$

$$= \phi r(\overline{\chi}_p,\overline{\chi}_s|\overline{\beta})\ d\overline{\chi}_s$$

where the "mixed" function ϕr plays the role of probability density for $\overline{\chi}_s$ and of probability function for $\overline{\chi}_p$ (see Wilks 1962, pp. 47-48).

The likelihood function is then defined as

$$L(\overline{\beta};\overline{\chi}_p,\overline{\chi}_s) = \phi r(\overline{\chi}_p,\overline{\chi}_s|\overline{\beta}) = P^{N-\overline{R}_p}(1 - P)^{\overline{R}_p}\ c \exp\{- \frac{1}{2\sigma^2} R_s\} ,$$

and the maximum-likelihood combined solution is that $\overline{\beta}$ maximizing L as a function of $\overline{\beta}$ for given $\overline{\chi}_p$ and $\overline{\chi}_s$.

L can be written in an alternate form which is illuminating;

$$L = c\ P^N \left(\frac{P}{1-P} \right)^{-\overline{R}_p} e^{- \frac{1}{2\sigma^2} R_s}$$

$$= k\ e^{-\left\{ \ln \left(\frac{P}{1-P} \right) \overline{R}_p + \frac{1}{2\sigma^2} R_s \right\}}$$

or

$$L(\overline{\beta};\overline{\chi}_p, \overline{\chi}_s) = k\ e^{-R_{ps}}$$

where R_{ps} is defined by

$$R_{ps} = a\overline{R}_p + bR_s \quad ,$$

with

$$a = \ln\left(\frac{P}{1-P}\right),$$

$$b = \frac{1}{2\sigma^2},$$

and k is a constant.

L is maximum for R_{ps} minimum. R_{ps} is denoted as the "combined P-S score" and is a linear combination of $\overline{R}_p$ and R_s. a and b play the role of "weights" of the P-wave and S-wave contributions to the combined solution and are dependent in an intuitively satisfying way upon the quality of the separate P-wave and S-wave solutions as measured by $P = \hat{P}$ and $\sigma = \hat{\sigma}$, respectively. $\hat{P}$ and $\hat{\sigma}$ are obtained either by the equations given above from separate P-wave and S-wave solutions done before the combined solution or, in some cases, from experience with previous comparable P-wave and S-wave solutions.

We have found that straightforward application of maximum-likelihood estimation leads to a solution having the intuitively desirable features which some investigators have sought to incorporate on an ad hoc basis. The choice of the measure of error is no longer arbitrary, and the correct procedure for "weighting" of P-wave and S-wave data is clearly indicated.

On the basis of the above formulations, which apply to any model of P-wave first motions and S-wave polarization angles, the P-, S-, and P-S maximum-likelihood solutions for the orientation of the focal planes are implemented by computing "score surfaces" or "error surfaces" for each type of data and model of source mechanism.

R_s, R_p, and R_{ps} are evaluated over a discrete grid of values covering the entire range of the parameters. The program for the S-wave data computes error surfaces for both type I (single-couple) and type II (double-couple) source models so that both may be compared with the P-wave error surface which is the same for type I or type II. The examples included here are for type II source models only. Because type I and type II P-wave and S-wave senses and directions of first motions are dependent only upon the orientation of the source mechanism, the source parameters consist of the three angles necessary to specify an arbitrary orientation of a coordinate triad in three dimensions.

The following explanation of the details of the solution procedure will facilitate the discussion of fiducial regions and is therefore given first.

E. Focal Plane Determination

Most work in determining the focal planes of an earthquake has been done using source models which are point sources. The observed data are plotted on a unit sphere, called the focal sphere, centered on the hypocenter. A unit position vector locates the observed data for station i on the focal sphere at the azimuth and dip along which the ray to station i leaves the source. Various projections have been used to display the focal sphere and the observed data in two dimensions. The techniques of solution and discussion of the results are dependent on the projection used. Stauder and others have made popular the use of the stereographic projection for the presentation of results. This projection, a conformal azimuthal projection, gives a nice representation of the focal sphere and makes the results fairly easy to visualize. Another projection also used (and used in this report) is the azimuthal equal-area projection--sometimes called the Schmidt net. This projection makes possible accurate comparisons of the relative sizes of contoured areas.

A plane through the center of the focal sphere intersects the focal sphere along the arc of a great circle path which projects into a circle on the stereographic projection and into a quartic curve closely approximating a circle on the azimuthal equal-area projection. The "pole" of such a plane is defined as the point where the unit vector normal to the plane intersects the focal sphere, or the projection of this point.

A discrete, incremental process is used in the program described in this report to compute the scores upon which the focal mechanism solutions and fiducial regions are based. The score is computed for an initial orientation of the nodal planes. The mechanism is then rotated incrementally through three independent angles until the focal sphere is covered.

Because of the radial symmetry of the radiation pattern of type I and type II source models, only one-half of the focal sphere need actually be sampled. For this reason, the above-mentioned projections of the focal sphere are of the lower hemisphere only. Points on the upper hemisphere are first mapped into their diametrically opposite points on the lower hemisphere. This fact has some (generally insignificant) implications for the constructions described below.

The orientation corresponding to the best score is the maximum-likelihood solution. As mentioned above, the solution for P-wave data may consist of a region (several points of the discrete grid) rather than one point only. The score computed at each position of the focal planes is saved for use in determining fiducial regions by the method described below.

The rotations are defined in a right-handed coordinate system at the focus having positive axes $\bar{x}$ (north), $\bar{y}$ (east), and $\bar{z}$ (down). The azimuth ϕ of a vector is the angle measured clockwise from north to the projection of the vector onto the $\overline{xy}$-plane. The spherical colatitude θ of a vector is the angle from the $+\bar{z}$-axis to the vector. The direction cosines of a unit vector X in the $\overline{xyz}$-system are given in terms of θ and ϕ by

$$X = \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} = \begin{bmatrix} \sin\theta \cos\phi \\ \sin\theta \sin\phi \\ \cos\theta \end{bmatrix} .$$

A plane is uniquely specified by its unit normal vector whose intersection with the unit sphere is the "pole" of the plane. If the unit normal to a plane Pl is given in terms of θ and ϕ, as above, then the strike and dip of plane Pl are $(\phi \pm \frac{\pi}{2})$ and θ, respectively. Here, "strike" denotes the azimuth of the line of intersection of plane Pl and plane $\overline{xy}$ (the horizontal plane), and "dip" denotes the dihedral angle ($<\pi/2$) formed by the Pl-plane and the $\overline{xy}$-plane.

The rotations are carried out by the equation

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\lambda & \sin\lambda \\ 0 & -\sin\lambda & \cos\lambda \end{bmatrix} \begin{bmatrix} \sin\theta & 0 & \cos\theta \\ 0 & 1 & 0 \\ -\cos\theta & 0 & \sin\theta \end{bmatrix} \begin{bmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \bar{x} \\ \bar{y} \\ \bar{z} \end{bmatrix},$$

or $X = D\bar{X}$ where $D = CBA$ and C,B,A are the indicated orthogonal 3 by 3 matrices representing rotations about current x-,y-,z-axes, respectively. Specifically, A represents a rotation of the coordinate axes about $\bar{z}$ through the angle ϕ measured from the $\bar{x}$-axis toward the $\bar{y}$-axis, B is a rotation about the current y-axis, through the angle $(\frac{\pi}{2}-\theta)$ measured from the x-axis toward the z-axis, and C is a rotation about the current x-axis through an angle λ measured from the y-axis toward the z-axis. $\bar{X} = (\bar{x},\bar{y},\bar{z})^t$ is a (unit) position vector of a point on the focal sphere in the above defined, earth-oriented coordinate system; and $X = (x,y,z)^t$ gives the coordinates of this same point in the system of the rotated focal mechanism.

Because of the orthogonality of D, the transformation from the rotated coordinates X to earth-fixed coordinates $\bar{X}$ is given by $\bar{X} = D^t X$. Also note that the rows of D are the direction cosines of the rotated axes. Thus,

$$D = \begin{bmatrix} \alpha_x & \beta_x & \gamma_x \\ \alpha_y & \beta_y & \gamma_y \\ \alpha_z & \beta_z & \gamma_z \end{bmatrix} ,$$

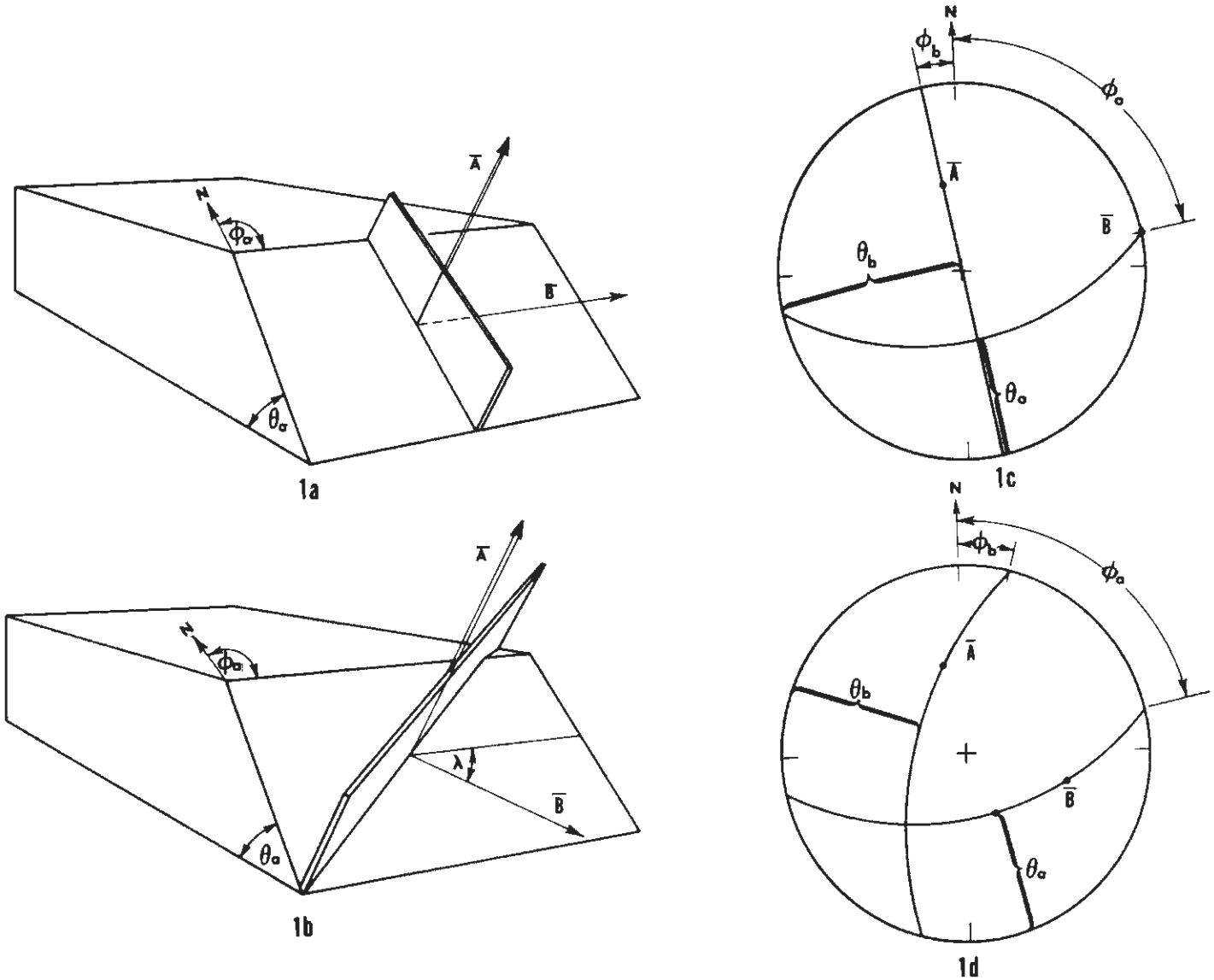

Figure 1.--Perspective view and corresponding lower hemisphere representation of the focal planes.

with an obvious notation for the direction cosines. Multiplying, one finds that α_x, β_x, γ_x are given in terms of θ and ϕ by the same equations as $X = (\alpha,\beta,\gamma)^t$ above. Therefore, θ and ϕ are the dip and azimuth of the (rotated) x-axis, the pole of the yz-plane, and this pole is not moved by the C-rotation through λ.

For convenience, denote the yz-plane as the "A-plane" and its associated pole (the x-axis) as the "A-pole". Similarly, denote the xz-plane as the "B-plane" and the y-axis as the "B-pole" (see figure 1). The designations "A-plane and B-plane" are a convenience for the purpose of describing the numerical solution only and do not imply any physical distinction between the two planes. The A- and B-planes are the nodal planes of P-wave and S-wave radiation patterns for a type II source that consists of two couples having equal forces and moments lying in the xy-plane and aligned with the x- and y-axes. Because of the orthogonality of the coordinate axes, the A- and B-planes are perpendicular and the A-pole lies in the B-plane and the B-pole lies in the A-plane.

For $\phi = \theta = \lambda = 0$, the initial orientation of the movable axes is x down, y east, and z south. ϕ is then incremented from 0 to 2π. For each value of ϕ, the incrementation of θ from 0 to π causes the A-pole to progress from downward to horizontal. Thus, the A-pole occupies successively the points of a grid covering the lower half of the focal sphere. Finally, for each pair of values ϕ and θ, that is, for each position of the A-pole, λ is incremented from 0 to π. Thus, the C-rotation through λ rotates the B-plane about pole A, and the B-pole rotates in the fixed A-plane while remaining on the lower half of the focal sphere.

To help illustrate this process, figure 1 shows perspective drawings of two orthogonal nodal planes and their corresponding representation on a stereographic projection before and after a λ-rotation. In figure 1a, the A-plane is the angular face of the block with normal vector $\bar{A}$. The B-plane contains $\bar{A}$ and has normal vector $\bar{B}$. ϕ_a and ϕ_b are the strikes of the A- and B-planes and θ_a and θ_b are their dips. (Figures 1a and 1b do not show ϕ_b or θ_b due to the difficulty of showing these angles on this figure.) In figure 1b, the B-plane has been rotated by the angle λ while remaining orthogonal to the A-plane, that is, $\bar{A}\cdot\bar{B} = 0$. In figures 1c and 1d, $\bar{A}$ is fixed while $\bar{B}$ moves along the trace of the A-plane during the λ-rotation.

When the A-plane has been incremented through all of its orientations, a grid of minimum error values will have been generated over the lower hemisphere of the focal sphere. These values can be contoured to show the shape of the minimum error surface, the most important features of which are the minima and the 95-percent fiducial regions.

If the grid spacing were indefinitely small, this procedure would result in a twofold redundant covering of the lower hemisphere. Each pair of positions of the two poles would be occupied twice, the second time with $\overline{A}$ and $\overline{B}$ interchanged. However, because of the finite grid spacing (usually 5°), the duplication is not exact. Because of the convenience in the contouring program of a complete uniform grid in θ and ϕ of the projected scores (the terminology is explained below in the discussion of fiducial regions) and of the desirability of a smaller increment size on λ, this redundancy is retained.

The λ-rotation is preferably incremented by a smaller step size because of the importance for good contours of defining an optimum score over the "hidden" angle λ. This is done by the grid search of λ alone without the aid of interpolation. The variation of the score with λ is separately plotted as in figure 12 only at the solution, if at all. On the other hand, the marginal scores on the grid of "explicit" angles θ and ϕ are used to construct the contoured regions such as those in figures 9, 10, and 11. Thus, the detailed variation of scores with θ and ϕ is interpolated and explicitly displayed, whereas the variation with λ does not have these advantages. Experience with the program and comparisons with exact regions constructed by the procedure described below have confirmed the desirability of a finer sampling of λ than of θ and ϕ to establish precisely the score projections before the θ, ϕ contouring is done. More complicated solution and contouring procedures have been rejected in favor of a simple, direct construction.

For each position of the A-pole, the best value of R_p and R_s (P-score and S-score) found during the rotation of the B-plane is assigned to the position of the A-pole on the projection of the lower half of the focal sphere. These values are contoured to produce nonlinear 95-percent fiducial regions about the P-wave solution and S-wave solution as explained below. At each step in the λ-rotation, R_p and R_s are stored for use in the combined solution. In the combined solution, the smallest combined score R_{ps} over λ is associated with each A-pole position. These values of R_{ps} are then contoured at equal-score intervals to show the likelihood surface for the combined solution. The minima of this surface give the maximum-likelihood solution for the position of the poles A and B. The results of these plots are shown in figure 11.

From Udias (1964), we have the equations of the computed polarization angles ε for both the type II (double-couple) and

the type I (single-couple) polarization angles. The equation for the double-couple polarization angles is

$$\tan\varepsilon = \frac{SH}{SV} = \frac{(\beta_x\, y + \beta_y\, x)\bar{x} - (\alpha_x\, y + \alpha_y\, x)\bar{y}}{2xy\bar{z} - (\gamma_x\, y + \gamma_y\, x)}.$$

The observed polarization angles are plotted on a stereographic projection as in figure 6. The theoretical values of polarization angles for each solution are computed from the above equation and plotted as shown in figure 7. This plot allows visual comparison with the observed values and is helpful for evaluation of a particular station's agreement with the solution.

The equation for the single-couple polarization angles is

$$\tan\varepsilon = \frac{SH}{SV} = \frac{-(\alpha_x\, \bar{y} - \beta_x\, \bar{x})\cos i_o}{(\alpha_x\, \bar{x} + \beta_x\, \bar{y})\cos^2 i_o - \gamma_x\, \bar{z}\sin^2 i_o}.$$

In our procedure, it is simple to replace one equation with the other in the computer program. However, experience has shown that the double-couple model with its larger number of adjustable parameters will usually fit the observed data better than the single-couple equation. For this reason and because Udias and Baumann (1969), Stauder (June 1968, December 1968), and others find a double-couple model to be acceptable in a majority of cases, we have used the type II solution in most work so far. The computer time required makes it undesirable to run every earthquake using both models. In the future, we hope to study some shallow earthquakes using the single-couple model.

The two orthogonal planes A and B define four quadrants of the focal sphere. The double-couple and single-couple models both predict the same first motions that are alternately compressional and dilatational in these quadrants. The observed or computed first motion χ_i at station i was defined as +1 for compression and -1 for dilatation. The sign of a quadrant is defined as the sign of the product xy, of the x- and y-coordinates in the rotated system, of any position vector $X = (x,y,z)^t$ in that quadrant. From the above equations $x = \alpha_x\, \bar{x} + \beta_x\, \bar{y} + \gamma_x\, \bar{z}$, $y = \alpha_y\, \bar{x} + \beta_y\, \bar{y} + \gamma_y\, \bar{z}$ where $\bar{X} = (\bar{x},\bar{y},\bar{z})^t$ is the position vector in the original, unrotated system. Thus, if χ_i^c is the computed first motion, $\chi_i^c = \mathrm{sgn}(x_i y_i)$, giving +1 in the first and third quadrants and -1 in the second and fourth quadrants in the xy-plane. This sign convention agrees with that of Stauder (1960).

An observation agrees with the theoretically computed first motion if the sign of the observation agrees with the sign of the quadrant for a particular orientation of the focal planes. Denoting the observed motion by χ_i^o, the P-wave score, R_p = number of agreements, can be evaluated from

$$R_p = \sum_{i=1}^{N} \{\frac{1}{2} + \frac{1}{2} \operatorname{sign}(x_i y_i \chi_i^o)\} \quad .$$

Here N is the total number of first motion observations and x_i, y_i are the coordinates of the ith station in the focal system, as above.

If the score computed in this manner is less than N/2, a reversal of the direction of one of the poles (moving it to the opposite point on the upper hemisphere) changes the sign of all quadrants and thus produces a new score of $N - R_p$ which is greater than $\frac{N}{2}$. The orientation of the planes is not affected, but the quadrant between the poles on the lower hemisphere is now dilatational rather than compressional. Thus, if $R_p < \frac{N}{2}$, then R_p is replaced by $N - R_p$.

III. EXACT P-WAVE LIMITS

The fiducial regions contoured by our programs will be still more meaningful if we consider an alternative method of finding P-wave focal plane solutions. The method described below will determine the solution and define the exact boundaries of the fiducial regions, or equal-score contours, for the poles of the solution. Given a unit vector X_i which defines the position of a station on the focal sphere, the plane whose normal is X_i intersects the focal sphere along a great circle which we will call the complementary circle of X_i. Two trial poles, for example, the positive x- and y-axes in the focal system, define two planes and corresponding complementary circles which divide the focal sphere into four quadrants. A lower hemisphere plot of these planes is shown by dashed lines in figure 2. If B_x and B_y denote the unit vectors along the x- and y-axes, the sign of the product $(A \cdot B_x)(A \cdot B_y)$ defines the sign of the quadrant containing the vector A.

The complementary circle of X_i divides the sphere into two hemispheres. If a vector B lies in the hemisphere containing X_i, then $X_i \cdot B$ is positive. On the complementary circle, $X_i \cdot B = 0$.

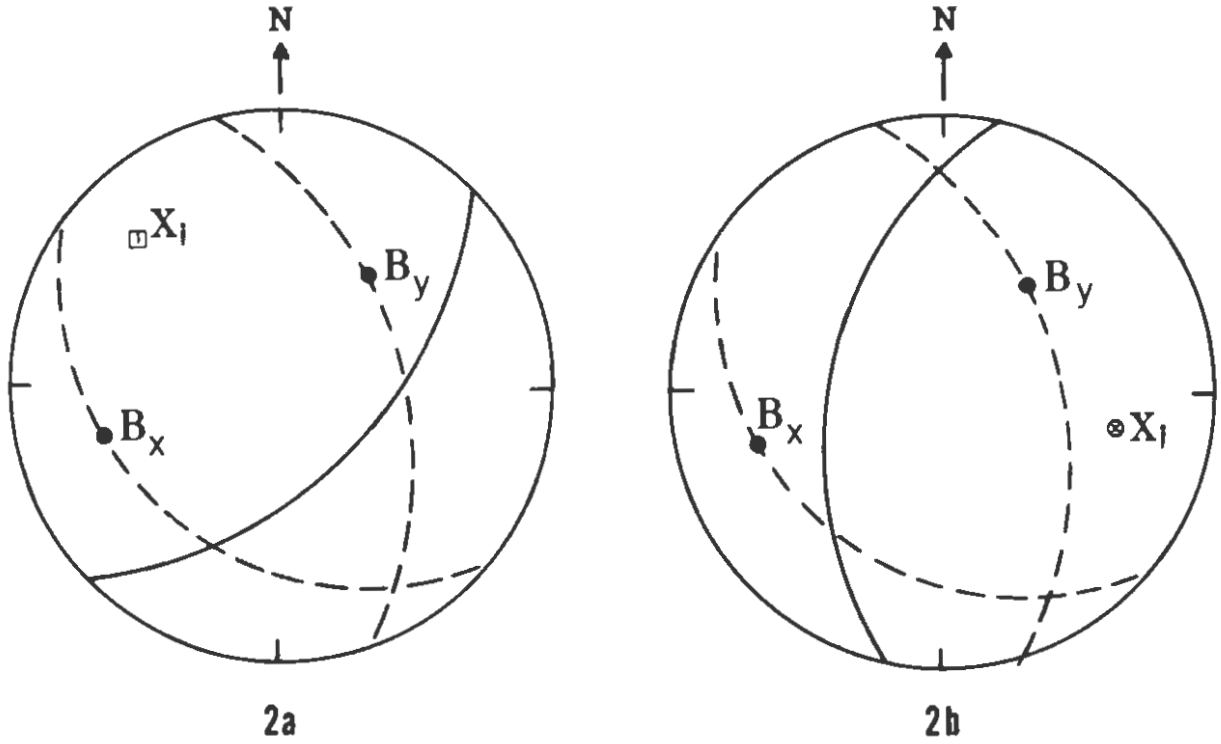

Figure 2.--Relation between the station complementary circles and the poles of the planes which define the compressional and dilatational quadrants.

When B lies in the hemisphere not containing X_i, $X_i \cdot B$ is then negative. Now if both trial poles B_x and B_y lie in the same hemisphere as in figure 2a, the product $(X_i \cdot B_x)(X_i \cdot B_y)$ will be positive and therefore X_i is in a positive quadrant. If the trial poles lie in opposite hemispheres as in figure 2b, then X_i lies in a negative quadrant and the poles are separated by the complementary circle of X_i.

Hence, we can conclude that counting the number of station vectors lying in negative quadrants is equivalent to counting the number of station complementary circles (SCC's) that divide the trial poles. Conversely, counting the number of stations in positive quadrants is equivalent to counting the number of SCC's that do not divide the trial poles. Note that "divide the poles" means the same as to cross the shortest great circle path joining the two poles. In fact, an arbitrary path between the poles crosses each SCC whose station vector lies in a negative quadrant an odd number of times and crosses each SCC whose station vector lies in a positive quadrant an even number of times.

The SCC's are differentiated in some fashion according to whether the observation at the station is compressional or dilatational. For example, the SCC's corresponding to compressional observations are plotted as solid lines and those corresponding to dilatational observations are plotted as dashed lines. Having selected any two orthogonal trial poles, the total number of agreements (i.e., the P-score R_p) is equal to the number of negative SCC's dividing the poles, $C_{(-)}$, plus the number of positive SCC's not dividing the poles. The last number is the total number of positive SCC's $N_{(+)}$, minus the number of positive SCC's dividing the poles $C_{(+)}$, giving $R_p = C_{(-)} - C_{(+)} + N_{(+)}$.

It may happen that one of the trial poles falls on the upper hemisphere of the focal sphere. In this case, a change of sign of the pole on the upper hemisphere transfers it to the diametrically opposite point on the lower hemisphere; this has the effect of changing the signs of all quadrants while leaving the planes in the same orientation. Thus, the score is changed to $N - R_p$. Because $N = N_{(+)} + N_{(-)}$, this gives $N - R_p = C_{(+)} - C_{(-)} + N_{(-)}$. By reversing the argument, we see that only trial poles and SCC's on the lower hemisphere need be considered if the score is computed by the equation

$$R_p = \begin{cases} T = C_a - C_b + N_b, & T \geq \frac{N}{2} \\ N = T, & T < \frac{N}{2} \end{cases} .$$

Here C_a is the number of SCC's of one type dividing the poles, C_b is the number of SCC's of the opposite type dividing the poles, and N_b is a constant equal to the total number of observations of the opposite type.

If C_a is the number of positive SCC's, then when $T<\frac{N}{2}$, the two poles on the lower hemisphere are separated by a dilatational quadrant; whereas, if $T\geq\frac{N}{2}$, they are separated by a compressional quadrant. This is mentioned as an aside because the sign of the quadrants is usually obvious from inspection of the observations. The case of both poles on the upper hemisphere is of no consequence because reversal of both poles leaves planes, signs, and scores unchanged. The maximum-likelihood solution consists of the region generated by all of those pairs of poles separated by the maximum possible number of SCC's of one type and the minimum possible number of the other type.

The preceding procedure applies whether or not the poles are orthogonal. However, the poles are required to be orthogonal in all solutions we have performed based on the incremental search described above.

With the use of this procedure, the exact boundary of the P-wave solution region can be obtained graphically. Consider a solution defined by five SCC's with compressions and dilatations plotted as solid and dashed lines, respectively, as in figure 3. A pole in the region G (dotted) and a pole in the region H (crosshatched) are separated by the maximum number (one) of dashed lines and the minimum number (zero) of solid lines. The requirement that the poles be orthogonal complicates the construction slightly. All poles in G or H may not possess an orthogonal pole in the other region. In the case represented in figure 3, H is large and the region consisting of all points orthogonal to points in H is also large and completely covers G. The reverse is not true. We can find the region of points orthogonal to any given region by drawing the complementary circles of the corner points of that region, as shown by the dash-dot lines in figure 3 for the region G. These lines will be great circles through pairs of stations. The area between these complementary circles contains all points orthogonal to points in G. Thus, considering both the orthogonality criterion and the station distribution, we find that the solution is bounded by the limits of the area K shown as dotted and crosshatched. In general, the region of points orthogonal to K may not cover all of the initial region G and this must be checked. In the example shown, G and K are orthogonal regions, that is, any point of either region has orthogonal points in the opposite region. Thus, G and K represent the regions for the poles giving the maximum-likelihood solutions for this example.

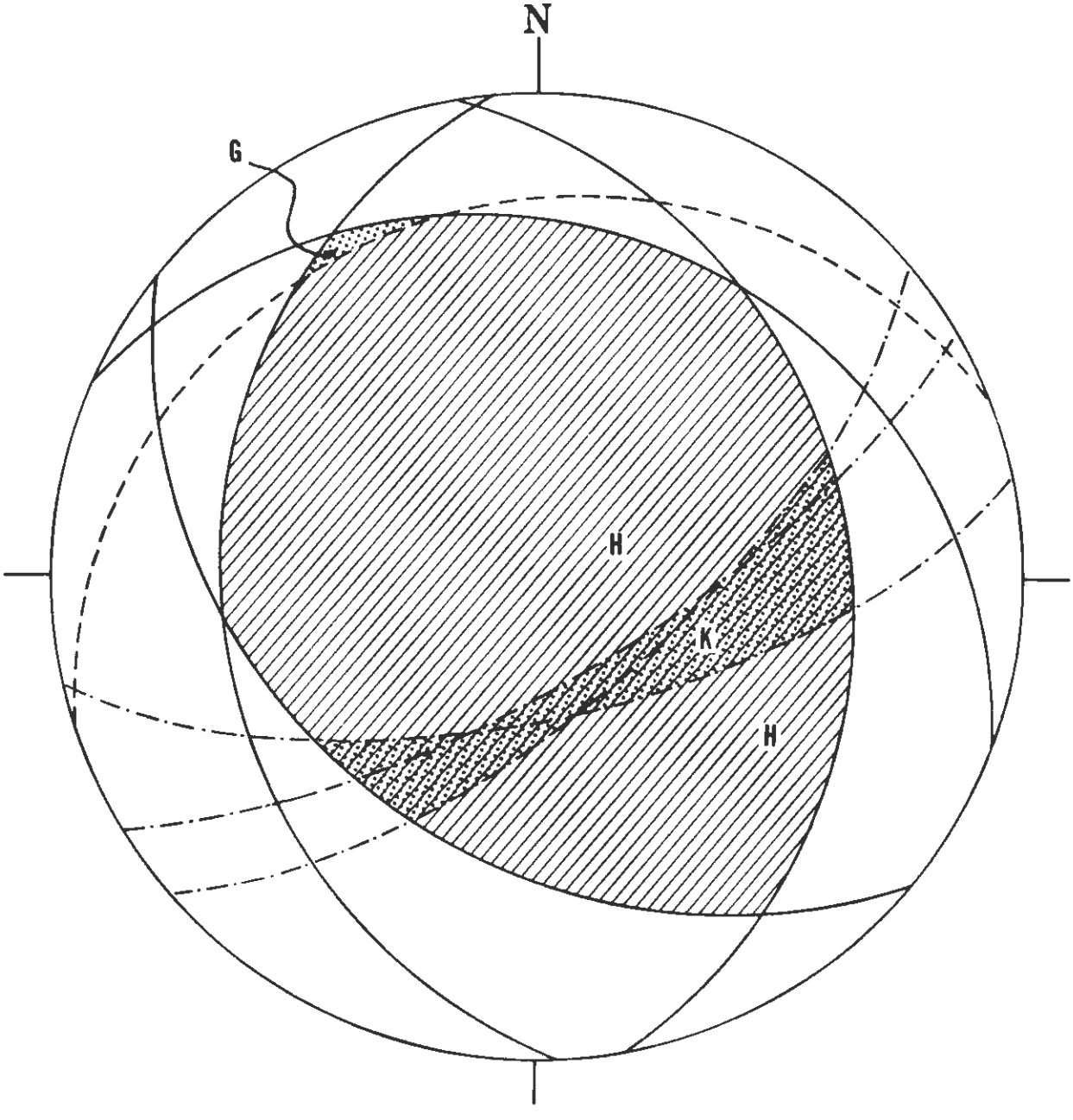

Figure 3.-- Construction of the exact boundaries for the maximum-likelihood P-wave solution.

Regions determined in this manner give the exact boundaries of the maximum-likelihood solution region. If we wish to find fiducial regions which will allow one or more additional stations to be in error, we can do so. To allow one more station to be wrong, we must discard one of the stations on the boundary of the maximum-likelihood region and see how that adds to the acceptable region. This process must be repeated with each station in turn on the boundary of the solution, mapping each time the additional area included. By discarding two or more stations at a time, this procedure can be carried to any fiducial level although it would become rather laborious to map the region graphically.

The program previously described was used to generate the solution and the illustrations shown in figures 5 through 7 and 9 through 12. This program samples the focal sphere incrementally. The sampled points were then contoured to show the approximate boundaries for the solution. It is interesting to compare the exact boundaries shown in figure 3 with the computer-contoured boundaries shown in figure 9. To make this comparison the example in figure 3 was constructed from five stations from the May 14, 1971, earthquake which defined the boundary of the maximum score region. The inner line of the computer-contoured plot should be equivalent to the exact boundaries. The two solutions cover the same general area. Considering that the points contoured by computer were generated on a five-degree grid, the agreement shown in figure 4 is reasonably good.

IV. S-WAVE AND P-WAVE FIDUCIAL REGIONS

A. Theory of the Fiducial Regions

Broadly speaking, a fiducial region consists of a standardized range of reasonable solutions. Standardization allows comparisons of the quality of different solutions by inspection of the relative size and shape of their associated fiducial regions. The determination of a fiducial region on more than one variable can be conveniently broken down into three subproblems, however, not without some overlap and not necessarily in this sequence: (a) choice of a criteria to define the boundary of the region; (b) determination of the size of the region, that is, choice of a particular boundary from among those meeting the criteria in (a); and (c) interpretation of the meaning of the region so obtained.

Boundaries of fiducial regions are conventionallly required to be equal-likelihood contours. Thus, the likelihood associated with any area within the region is greater than that of an equal area lying entirely without. Because of the nonlinearity of the functions involved, the equal-likelihood contours are not simple, concentric geometric curves such as the familiar ellipses in linear least-squares, but are relatively complex functions of

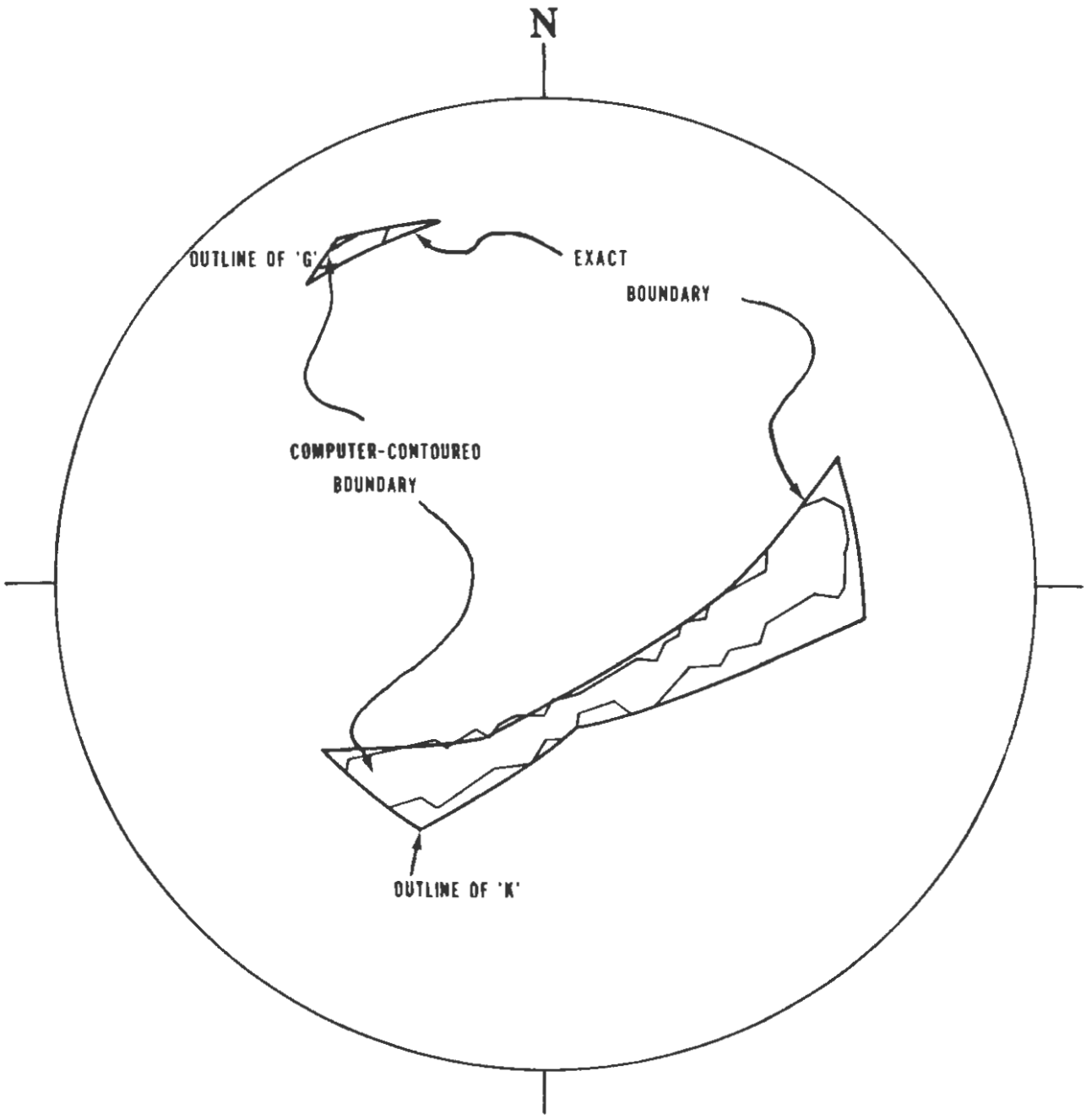

Figure 4.--A comparison between the computer-contoured and the exact boundaries for the maximum-likelihood P-wave solution.

the station distribution, the data, and the functional and statistical model assumed. Even in those nonlinear problems that still permit an exact analytical solution, analytical small-sample equal-likelihood contours are unlikely to be practically obtainable. The same applies, a fortiori, to subproblems (b) and (c).

There are two standard approaches which can overcome this difficulty in practice: (1) linearization of the model in the vicinity of the solution and use of the standard linear least-squares error-analysis procedures, or (2) numerical point-by-point construction of the fiducial regions. Even after the linearization in approach (1), the solution to subproblems (b) and (c) presents difficulties unless the joint distribution of error is multivariate normal. Under this assumption, complete solutions to subproblems (a), (b), and (c) are available for the linearized problem. In the nonlinear case only, a large-sample (large N) result is available for the exact distribution of maximum-likelihood estimators and consequent large-sample fiducial (confidence) regions. This result has only limited applicability in the present problem, but will be returned to briefly in connection with S-wave fiducial regions. The numerical construction, approach (2), has the advantage of producing the maximum-likelihood solution and the solutions to subproblems (a) and (b) by consistent procedures that use the same numerically generated values.

The boundaries for both S-wave and P-wave fiducial regions are contours of the score surfaces which are numerically generated as described above in connection with the maximum-likelihood solution procedure. While the determination of P-wave regions is entirely numerical, it is more convenient to determine the size of the S-wave fiducial regions using a result from linear least squares carried over as an approximation to the nonlinear case. Thus, P-wave data are handled entirely by approach (2); however, for computation of the S-wave fiducial regions, approach (1) is used to determine the size of the region and approach (2) to determine the boundary configuration. Inspection of the regions thus obtained (see figures 9, 10, and 11) confirms the significant departure of the contours from the elliptical shapes expected from a linear, or large-sample,model. That is, the nonlinearity of the model has a significant effect on the shape of the range of reasonable variation.

Before considering the statistical aspects of subproblems (b), content, and (c), interpretation, there are first some considerations that arise because the parameter space is three-dimension while the representation of regions is, of necessity, two-dimensional. The parameters are the three angles ϕ,θ,λ that specify the rotation carrying $\overline{X}$ into X.

In a multivariate problem, three types of fiducial regions are distinguished, just as three types of probability densities

(continuous variables) or functions (discrete variables) are distinguished: joint, marginal, and conditional. For the moment, we will postpone comment on any subtleties concerning the precise probability statements that can be made about fiducial regions. Roughly speaking, a joint 95-percent region is a volume in the curvilinear $\theta\phi\lambda$-space that has a 95-percent probability of covering the true values of θ,ϕ,λ. On the other hand, a marginal 95-percent region on θ and ϕ has a 95-percent probability of containing the true values of θ and ϕ, whatever the true value of λ. A conditional 95-percent region on θ and ϕ, conditional on λ_o, has a 95-percent chance of including the true values of θ and ϕ when λ is constrained to equal λ_o. While these distinctions may appear at first to be inconsequential, they are, in fact, central to the correct construction and interpretation of multivariate-confidence regions.

The only kind of region constructed here is a three-dimensional joint fiducial region in θ,ϕ, and λ. This three-mensional region is then visualized by means of its projections onto an azimuthal equal-area map of the lower focal hemisphere, just as a volume in Cartesian x-, y-, z-coordinates can be visualized by its projection upon the coordinate planes. The $\theta\phi\lambda$-space itself is curvilinear and cyclic, not Cartesian, and is therefore difficult to visualize as a whole. However, the basic idea of projection along tangential coordinate curves holds.

This tangential projection is effected numerically by contouring scores that are minimal over λ on a $\theta\phi$-grid. In the $\theta\phi\lambda$-space, such a minimum Rm (or maximum, depending on the score in question) is a point of tangency of a $\theta\phi$-constant, λ-variable coordinate line, with the three-dimensional, equal-score surface on which the score is Rm. A contour line of constant score on the $\theta\phi$-surface is generated by all those λ-coordinate lines tangent to the equal-score surface.

The projection of the lower hemisphere thus contains two two-dimensional "views" or curvilinear projections of the three-dimensional surface bounding the fiducial region and embedded in the $\theta\phi\lambda$-space. A third "view" could be obtained by contouring scores about the third pole, or any other direction fixed in the X-system. Only the A- and B-pole contours are shown, however, because these are most directly related to the position of the two nodal planes and are sufficient to describe the uncertainty in their orientation.

To make possible the comparison of different fiducial regions, it is necessary that the regions be standardized so that their content is in some sense the same. Recall that subproblem (c) involves the definition and interpretation of a suitable measure of content and that subproblem (b) deals with the procedure for finding a particular likelihood contour which will realize a certain content chosen in advance. It is generally acknowledged

that the ideal measure of content in this situation is probability. Thus, an ideal measure of the range of variability is a region constructed about the particular solution (an estimate of the unknown, true parameters), with the property that the probability of the region including the true parameters is a constant (say 95-percent) set in advance.

Such a region, if it can be found, is called a "strict confidence region." In a sense, all fiducial regions are approximations to strict confidence regions. Such regions may still differ among themselves, however, because of necessary compromises with the exact theory or because of differing interpretations of the acceptable meanings of "probability."

There is one method for construction of strict confidence regions that uses only the joint probability function $Pr(\overline{\chi};\overline{\beta})$ of data $\overline{\chi}$, dependent on fixed but unknown parameters $\overline{\beta}$. No additional new probability functions are introduced except those derivable by straightforward summation (Rao 1965, and Cramér 1946). In the following description of this construction, it is assumed that both the observation and parameter spaces are discrete. Similar equations apply when both are continuous. In the discrete case, an equation such as $\sum_A Pr = 1 - \alpha$ is understood to mean that $\sum_A Pr$ is the smallest sum $\geq 1 - \alpha$. Exact equality is not usually possible because the sum increases by discrete jumps.

The probability function $Pr(\overline{\chi};\overline{\beta})$ gives the probability of $\overline{\chi}$ for a particular value of $\overline{\beta}$. On the basis of these data $\overline{\chi}$, the maximum-likelihood solution is $\hat{\beta}(\overline{\chi})$. Then $Pr(\hat{\beta};\overline{\beta}) = \Sigma Pr(\overline{\chi};\overline{\beta})$ where the sum is over all $\overline{\chi}$ that produce $\hat{\beta}$ as the maximum-likelihood estimate. Thus, for a given $\overline{\beta}$, we can associate with every point in the parameter space the probability of finding that point $\hat{\beta}$ as the maximum-likelihood solution $Pr(\hat{\beta};\overline{\beta})$. If each $\overline{\chi}$ determines a unique $\hat{\beta}$, then $\sum_{all\,\hat{\beta}} Pr(\hat{\beta};\overline{\beta}) = 1$ because $\sum_{all\,\overline{\chi}} Pr(\overline{\chi};\overline{\beta}) = 1$. A region $A(\overline{\beta})$ about $\overline{\beta}$ so that $\sum_A Pr(\hat{\beta};\overline{\beta}) = 1 - \alpha$ for a given $1 - \alpha$, say 95-percent, is called, for reasons not gone into here, an "acceptance region." Some criteria, such as $Pr(\hat{\beta};\overline{\beta})$ = constant, must be established for the boundary of the acceptance region. Thus, with probability $1 - \alpha$, $A(\overline{\beta})$ contains $\hat{\beta}$ (the maximum-likelihood estimate) when the true value of the vector of the parameters is $\overline{\beta}$. Now let $C(\hat{\beta}_1)$ be the set of all $\overline{\beta}$ such that their corresponding $A(\overline{\beta})$'s all contain a particular estimate $\hat{\beta}_1$. Then C is a strict $1 - \alpha$ confidence region about $\hat{\beta}_1$; that is, the probability that the unknown true value of $\overline{\beta}$ is covered by C is $1 - \alpha$. For, by construction, C covers $\overline{\beta}$ if and only if $A(\overline{\beta})$ covers $\hat{\beta}_1$; that is, the two events $\{\overline{\beta}\varepsilon C(\hat{\beta}_1)\}$ and $\{\hat{\beta}_1\varepsilon A(\overline{\beta})\}$ are equivalent and therefore have the same probability

$1 - \alpha$. In contrast to the one-dimensional case, the multivariate case involves some arbitrariness arising from the choice of other criteria for the boundary of $A(\bar{\beta})$, $C(\hat{\beta})$, or both.

This definition of a strict-confidence region, while logically sound, has been criticized as being nonintuitive. More to the point are the criticisms that: (1) the properties of linearity, continuity, and homogeneity of the A's, of symmetry of $\phi(\hat{\beta};\bar{\beta})$, and so forth, that make an analytical construction of a strict confidence region possible in some cases simply do not apply in the case of S-wave and P-wave data; (2) direct numerical construction of the grid of acceptance regions is out of the question because of the magnitude of the computations involved; and (3) the correct extension of this derivation for strict confidence regions to mixed discrete and continuous Pr's and to set estimators ($\hat{\beta}$ not uniquely defined) is problematic. At present no practical procedure is known for construction of strict confidence regions by this method in the problem at hand.

In some problems, it may be expedient to use as an approximate fiducial region the acceptance region constructed about the particular solution $\hat{\beta}$, taken as true. The correct probability statement is: Under the hypothesis that $\bar{\beta}$, the true value, is in fact equal to $\hat{\beta}$, the particular solution obtained, the probability that another solution $\hat{\beta}$ will fall within R is 95 percent. This region gives an idea of the scatter introduced into the solution process by random errors in data. The construction of this single acceptance region still involves difficulties in the general case.

A second method for construction of strict confidence regions requires a statistic $s(\bar{\chi},\bar{\beta})$ which is a function of the parameters $\bar{\beta}$, but whose probability function Pr(s) is not. Then a probability statement about s can be made without knowledge of the parameters and can sometimes be converted into a probability statement about the parameters. For example, if $Pr\{s(\bar{\chi},\bar{\beta})<s_1\} = 1 - \alpha$ and if $\{s(\bar{\chi},\bar{\beta})<s_1\}$ iff $\{\bar{\beta}\epsilon C(\bar{\chi})\}$, then $C(\bar{\chi})$ is a $(1 - \bar{\alpha})$-confidence region for $\bar{\beta}$ based on the observations $\bar{\chi}$. The linear-normal reggions described below in connection with S-wave data are most easily derived from this viewpoint, although an exact extension of this method to nonlinear, non-normal data is again not known.

There are three general constructions that are essentially different from the above two methods and that avoid some of their difficulties. The differences between the three lie more in the interpretation than in computational detail. For the purposes of the discussion in this report, the three constructions have been lumped together as "Bayesian" because of (i) computational similarities; (ii) historical connections; (iii) the common feature that each construction involves the introduction of probability measures or conventional measures of content (such as a priori probability, fiducial probability, or percent likelihood) whose

significance must be postulated rather than deduced from that of $Pr(\bar{\chi};\bar{\beta})$; and (iv) the common features that only one point in the sample space of observables $\bar{\chi}$, that which is actually observed, enters into the equations. Thus, no summations of probability over $\bar{\chi}$, such as those needed to construct acceptance regions, are involved. The newer consensus seems to be that Bayesian procedures are simpler, more flexible, and more appealing to intuition than are strict, non-Bayesian constructions; in addition, they have the advantage of being possible.

The three "Bayesian" constructions are discussed below.

(1) Percent-likelihood region. Let $K = \sum_{\text{all}\,\bar{\beta}} L(\bar{\beta};\bar{\chi})$, then R bounded by L = constant and such that $\sum_R (L/K) = 1 - \alpha$ is a $100(1 - \alpha)$ percent-likelihood region. This is a minimal interpretation, stripped of all statistical implications, that is favored by physicists and others for just the sort of statistically intractable problem we have at hand. Nothing is asserted except that R contains, for example, 95 percent of the likelihood over $\bar{\beta}$ for a given $\bar{\chi}$. This interpretation achieves the goal of standardization of content.

(2) Fiducial region. There does not seem to be a clear exposition of fiducial probability in the multivariate case requiring renormalization. For present purposes, fiducial probability apparently amounts to nothing more or less than the assertion that because (renormalized) likelihood obeys the defining postulates for a probability function, and for other more obscure reasons not gone into here, renormalized likelihood is, in fact, a probability function denoted by the term "fiducial probability" to distinguish it from conventional probability. The region R constructed exactly as in construction (1) is called a 95-percent fiducial region.

(3) Classical Bayesian confidence region. This is a region R, given by $\sum_R Pr(\bar{\beta}|\bar{\chi}) = 1 - \alpha$, Pr = constant on boundary, where $Pr(\bar{\beta}|\bar{\chi})$ is the a posteriori probability function of $\bar{\beta}$ conditional upon $\bar{\chi}$, obtained from Bayes' formula

$$Pr\ (\bar{\beta}|\bar{\chi}) = L(\bar{\beta};\bar{\chi})\ Pr(\bar{\beta})/T.$$

Here $T = \sum_{\text{all}\,\bar{\beta}} L(\bar{\beta};\bar{\chi})Pr(\bar{\beta})$ where $Pr(\bar{\beta})$ is the a priori probability function of $\bar{\beta}$.

The problem sometimes encountered is that there may be no empirical hold on $Pr(\bar{\beta})$ and that there may be more than one reasonable choice of parameters and associated a priori Pr that do not

lead to identical results but for which there is no clear rule for choosing one over another. The only situation considered here is that of complete a priori ignorance about $\bar{\beta}$. This is usually expressed by $Pr(\bar{\beta}) =$ where $\sum_{\bar{\beta}} \varepsilon = 1$ and ε is constant, although the case of mixed discrete observations and continuous parameters admits other possibilities. (Because the parameter space spanned by $\theta\phi\lambda$ is finite, we need not be concerned with the limit as $\varepsilon \to 0$, as is sometimes necessary.) In this case, $Pr(\bar{\beta}|\bar{\chi}) = L/K$ with $K = \sum_{all\bar{\beta}} L$ so that the classical Bayesian confidence regions are identical with those obtained from construction (1) or (2). The statement is allowed that, on the basis of the data $\bar{\chi}$, there is a 95-percent a posteriori probability that the true value of $\bar{\beta}$ lies in R.

Historically, constructions (1) and (2) are outgrowths of construction (3) that attempt to avoid the introduction of a priori probabilities while retaining the computational advantages of a Bayesian approach. In this report, the term "fiducial region" is used, not because of a special preference for construction (2) above, but as a neutral term which has been used by others to avoid the implication that regions are either strict confidence regions on the one hand or classical Bayesian confidence regions on the other. A fiducial region is defined by the specification of a computational procedure, the results of which may be interpreted as in construction (1), (2), or (3).

Although the demonstration is lengthy, all of the above constructions, non-Bayesian and Bayesian (understood to involve only the "complete-ignorance" uniform a priori probability as described above), can be shown to lead to the same regions when applied to linear-normal adjustments incorporating the standard assumptions. That is, the joint regions finally obtained will all be ellipsoids (for $\mu = 3$) of identical size and shape. Thus, in the linear-normal case, there are several simultaneously correct (or at least valid alternative) probability statements which can be made to aid the interpretation of the regions. In the absence of linearity, normal statistics, continuity, and large samples, no such unification has been found. Mixtures of discrete observables and continuous parameters give rise to still more alternatives and problems of interpretation.

This aforementioned list does not exhaust all possibilities for construction of confidence regions. The aim has been only to give a brief summary of methods available for the solution of subproblems (b) and (c) in the case of S-wave and P-wave first motion data.

B. S-Wave Fiducial Regions

If the S-wave model is linearized in the vicinity of $\hat{\beta}$, the following results hold (Hamilton 1964, chapter 4). If R_{LS} is the

minimum R_s at $\hat{\beta}$ where R_s is the S-wave score, defined as above, and R denotes R_s at some $\bar{\beta} \neq \hat{\beta}$, then $(R - R_{LS})/R_{LS}$ is distributed as $(b/\nu)F_{b,\nu}$ where b is the dimension of $\bar{\beta}$, (three), and ν is the "degree of freedom", ($\nu = n - u = N - 3$), associated with $\hat{\sigma}^2$. Then $R \leq \{1 + \frac{b}{\nu} F_{b,\nu,\ 1-\alpha}\} R_{LS}$ is a $(1 - \alpha)$-fiducial region on $\bar{\beta}$.

The use of exact nonlinear contours whose size is determined by the statistics of the linear case is recommended in Draper and Smith (1966). The proof that the region so constructed constitutes a fiducial region for $\bar{\beta}$ (while on the face of it, the probability statement refers to a function of R), rests in the linear case upon the identity; $(R - R_{LS})/R_{LS} = (\bar{\beta} - \hat{\beta})^t N(\bar{\beta} - \hat{\beta})/R_{LS} = (\bar{\beta} - \hat{\beta})^t \hat{M}_{\hat{\beta}}^{-1} (\bar{\beta} - \hat{\beta})/\nu$, where $\hat{M}_{\hat{\beta}} = N^{-1}\hat{\sigma}_o^2$ and $N = A^t WA$, the normal equation coefficient matrix formed from the observation equations, $\bar{V} = \{\bar{\varepsilon}(\bar{\beta}_o) - \bar{\varepsilon}^o\} + A(\bar{\beta} - \bar{\beta}_o)$, that result from Taylor's series linearization of $\bar{\varepsilon}(\bar{\beta})$, with $A = \left.\frac{\partial \bar{\varepsilon}}{\partial \bar{\beta}}\right|_{\bar{\beta}_o}$. The ellipsoid defined by $(\bar{\beta} - \hat{\beta})^t \hat{M}_{\hat{\beta}}^{-1} (\bar{\beta} - \hat{\beta}) \leq bF_{b,\nu,\ 1-\alpha}$ is the $(1 - \alpha)$-confidence region about $\hat{\beta}$ in the linear problem. If $\nu \to \infty$, that is, σ_o^2 is assumed to be known without error, the last equation becomes $(\bar{\beta} - \hat{\beta})^t M_{\hat{\beta}}^{-1} (\bar{\beta} - \hat{\beta}) \leq \chi^2_{b,1-\alpha}$ where $M_{\hat{\beta}} = N^{-1} \sigma_o^2$. In most linear problems, this ellipsoid is more easily constructed by diagonalization of N or M than by contouring of R-values.

The equivalence of the two procedures is assured in the linear case by the identity just mentioned. In the nonlinear case, the contouring of R-values gives the correct shape of the fiducial region; the use of the $R = \{1 + \frac{b}{\nu} F_{b,\nu,\ 1-\alpha}\} R_{LS}$ contour as a $1 - \alpha$ region is justified as an approximation by an argument using a Taylor's series expansion of R about $\hat{\beta}$. This expansion gives $R - R_{LS} = \frac{1}{2} (\bar{\beta} - \hat{\beta})^t Q(\bar{\beta} - \hat{\beta}) + \ldots$ where $Q = \left.\frac{\partial^2 R}{\partial \bar{\beta} \partial \bar{\beta}}\right|_{\hat{\beta}}$. There is no term linear in $(\bar{\beta} - \hat{\beta})$ because $\hat{\beta}$ is a stationary point of R. From $R = \bar{V}^t W \bar{V}$ and $\bar{V} = \bar{\varepsilon}(\bar{\beta}) - \bar{e}^o$, one finds, for W diagonal, $Q = 2\{n_{ij} + \sum_\ell w_\ell v_\ell \frac{\partial^2 \varepsilon_\ell}{\partial \beta_i \partial \beta_j}\}$ where $\{n_{ij}\} = N$ and $\{\varepsilon_i\} = \bar{\varepsilon}(\bar{\beta})$. Thus, $R - R_{LS} = (\bar{\beta} - \hat{\beta})^t N(\bar{\beta} - \hat{\beta})$ + terms that are

at least cubic in small quantities $(\overline{\beta} - \hat{\beta})$ and $\overline{V}$. Neglecting these higher terms leaves the identity from which the desired result follows just as in the truly linear case.

This approximation in the neighborhood of $\hat{\beta}$ also justifies the use of $\hat{\sigma}_o$ derived from the linear theory. $(\nu\hat{\sigma}_o^2/\sigma^2) = (R_s/\sigma^2)$ is distributed as χ_ν^2, a fact which can be used to construct a confidence interval for σ^2 about $\hat{\sigma}_o^2$ (Hamilton 1964).

Note that, because the expected value of v_i, $E(v_i)$ and, therefore, of any linear combination of v_i's is zero, $E\left\{\frac{1}{2}Q\right\} = N$. Define I, Fisher's information matrix, by $I = -E\left\{\frac{\partial^2 \ln L}{\partial\overline{\beta}\partial\overline{\beta}}\right\}$. Since $L(\overline{\beta}) = c \exp\left\{-\frac{1}{2\sigma^2} R_s(\overline{\beta})\right\}$, then $\frac{\partial^2 \ln L}{\partial\overline{\beta}\partial\overline{\beta}} = -\frac{1}{2\sigma^2}Q$ and $I = N\frac{1}{\sigma^2}$. It is known that for large numbers of observations, a maximum-likelihood estimate $\hat{\beta}$, based on likelihood L, is approximately (multivariate) normally distributed, with mean $\overline{\beta}$ and $M = I^{-1} = N^{-1}\sigma^2$. Thus, as the number of observations increases, the linear-normal theory becomes more accurate. Intuitively explained, this means that for large ν the fiducial region becomes smaller (generally N^{-1} behaves like $\frac{1}{\nu}I$) and in a small enough neighborhood of the solution the R = constant contours become ellipses.

The preceding has been in the nature of a summary of relevant points only, with no attempt at derivational completeness. The analysis of the linear-normal adjustment is well established and extensively discussed in the statistical literature (Hamilton 1964, and Graybill 1961). We have found the extension to the nonlinear case to be an eminently satisfactory and practical approximation that greatly simplifies the construction of S-wave fiducial regions.

C. P-Wave Fiducial Regions

As noted previously, the maximum-likelihood estimate based on P-wave first motion data is not unique, but consists of a region, possibly disconnected, in the parameter space. For given data $\overline{X}_p = \{\chi_i^o\}$, the three-dimensional parameter space can be subdivided into "constant-score" volumes in which R_p is constant. Constant score also means constant likelihood. Each of these volumes is imaged on the projection as two constant-score regions,

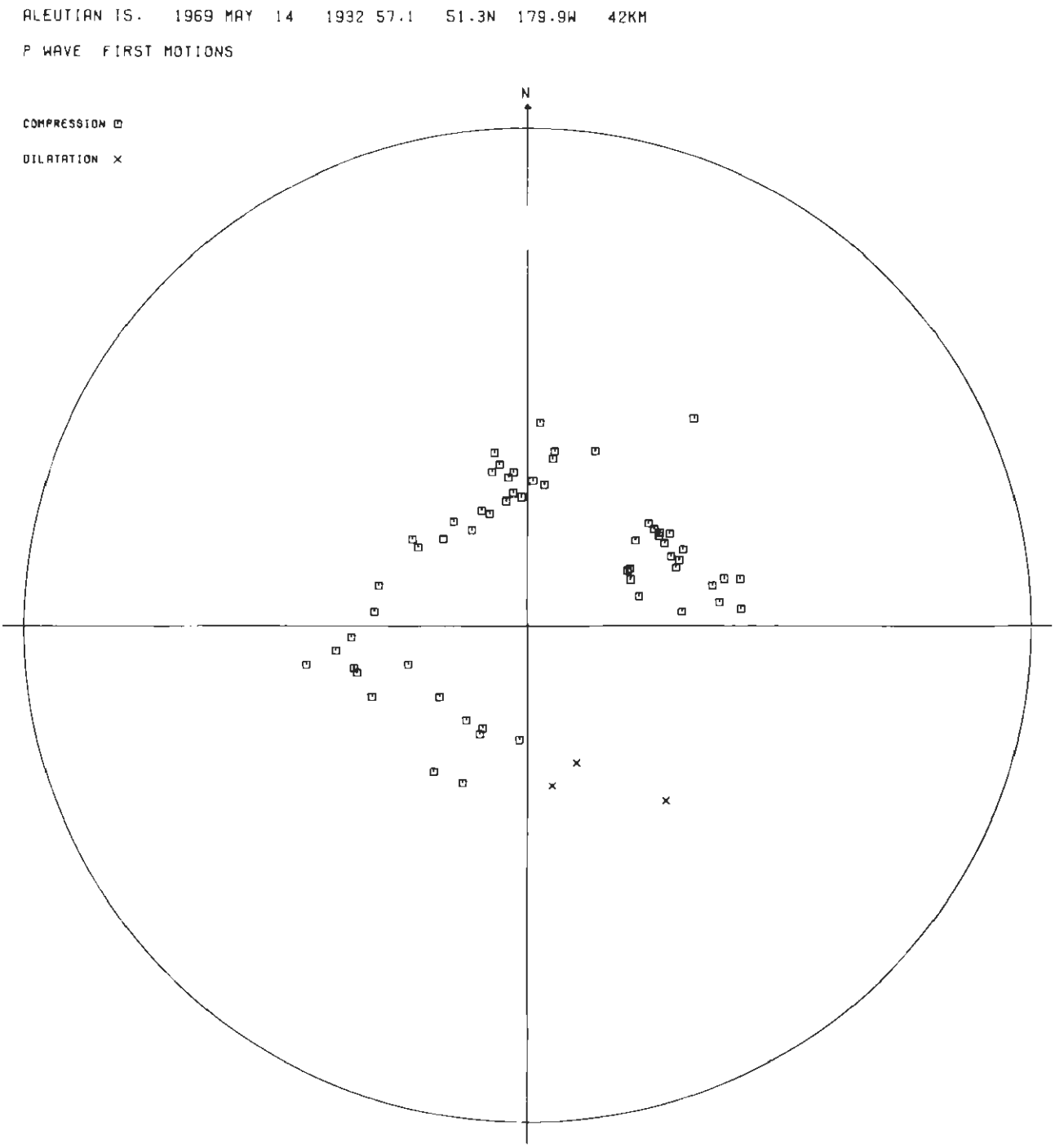

Figure 5.--Computer plot showing P-wave first motion data on the lower hemisphere.

one about each pole. In the following, "region" may refer either to a volume in the parameter space or to the areal projections of such a volume. The constant-score regions may be further subdivided into "constant-combination" regions (CCR's) in which the same R_p stations agree. Thus, in a CCR, the same specific stations (R_p in number) agree and the same $N - R_p$ stations disagree for any $\bar{\beta}$ in the region. If instead of $P_i = P$, the P_i are all different but constant (independent of $\bar{\beta}$), the CCR's are identical with the constant-score regions.

The CCR's represent the finest resolution of the parameter space that is possible on the basis of P-wave first motion data for a given station distribution on the focal sphere. No distinct points in the same CCR can be discriminated by any evidence from the P-wave first motion data. The partitioning of the sample space into CCR's is completely determined by the station distribution before the introduction of specific data $\bar{\chi}_i^o$. A CCR can be defined as the set of all $\bar{\beta}$ for which $\bar{\chi}_p^c(\bar{\beta}) = \{\chi_i^c\}$ is constant. Not all 2^N possible combinations of + 1's and - 1's in $\bar{\chi}_p^c(\bar{\beta})$ are produced by some $\bar{\beta}$. Therefore, for a given $\bar{\chi}_p$, not all $\binom{N}{R}$ possible combinations of agreements and disagreements producing a score of R are attained for some $\bar{\beta}$. Denote the number of combinations producing a score of R by C_{R_p} for later use. Constant-score regions are built up out of CCR's. The specific scores obtained and the configuration of the constant-score regions depend on $\bar{\chi}_p^o$, while the configuration of CCR's does not.

On the basis of data $\bar{\chi}_p^o$, each CCR has associated with it the likelihood $P^{R_p}(1 - P)^{N - R_p}$. That is, for $\bar{\beta}$ in a particular CCR denoted by A, the likelihood is $L(\bar{\beta}\varepsilon A;\bar{\chi}_p^o) = Pr(\bar{\chi}_p^o|\bar{\beta}\varepsilon A) = P^{R_p}(1 - P)^{N - R_p}$. Thus, the continuous parameter space has been divided into a finite number of discrete events ($\bar{\beta}\varepsilon A_j$), where $j = 1 \ldots m$ and m is the number of CCR's. The sum of likelihood over these events is given by

$$\sum_j L(\bar{\beta}\varepsilon A_j;\bar{\chi}_p^o) = \sum_j P^{R_p}(1 - P)^{N - R_p} = \sum_{R_p} \left\{ C_{R_p} P^{R_p}(1 - P)^{N - R_p} \right\}.$$

ALEUTIAN IS. 1969 MAY 14 1932 57.1 51.3N 179.9W 42KM

OBSERVED S-POLARIZATION ANGLES

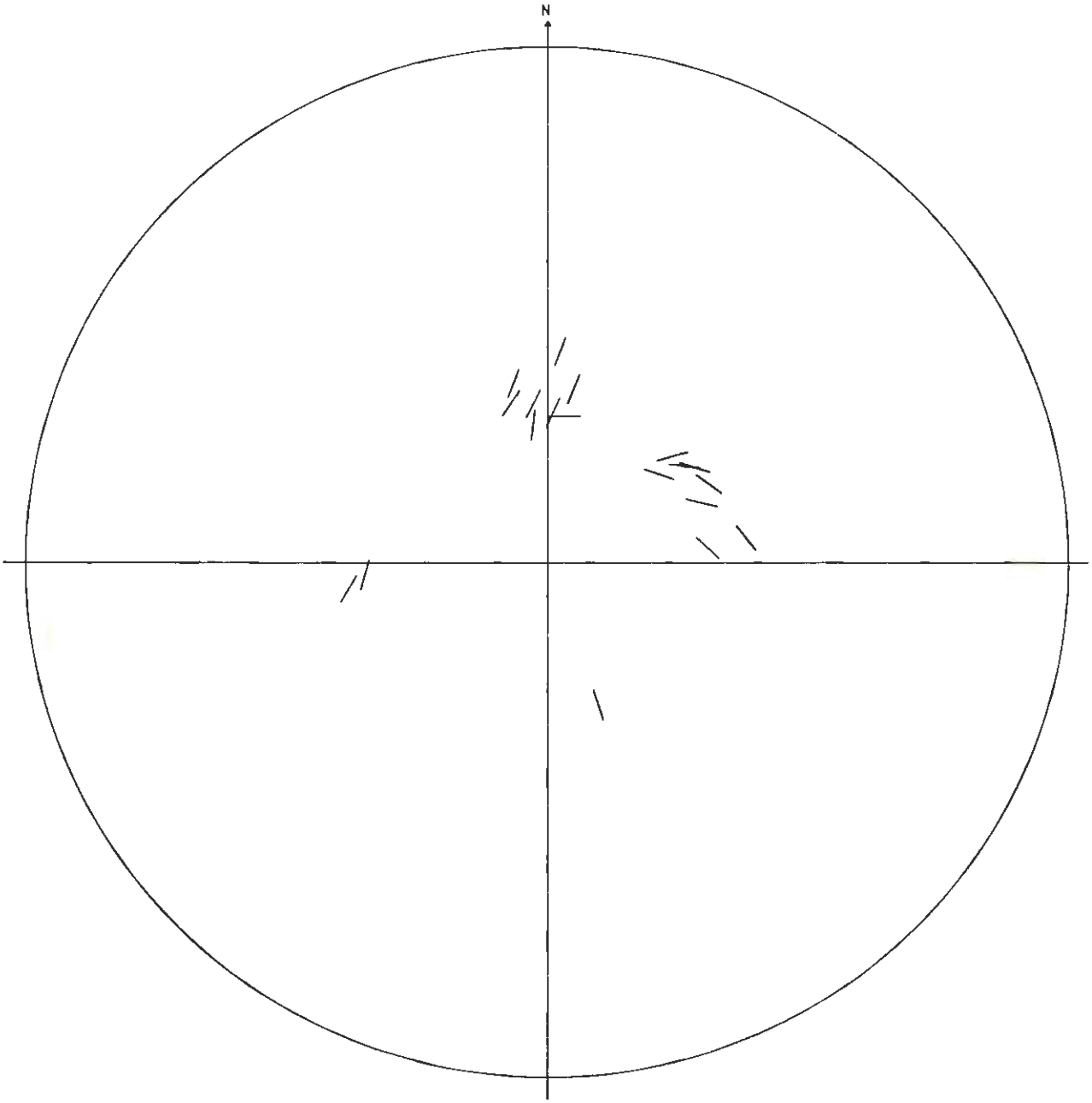

Figure 6.--Computer plot showing observed S-polarization angles on the lower hemisphere.

Let $K = \sum_{\text{all} R_p} \left\{ C_{R_p} P^{R_p} (1 - P)^{N - R_p} \right\}$; then region F, defined so that $\sum_F \left\{ C_{R_p} P^{R_p} (1 - P)^{N - R_p} / K \right\} = f$, is a 100f-percent fiducial region for $\bar{\beta}$ about $\hat{\beta}$ where f is the smallest value of the sum $\geq 100(1 - \alpha) = 95$ percent. The summation starts with R_{max} and proceeds downward, that is, $R_{max} - 1$, $R_{max} - 2$, and so forth. Region F is bounded by that equal-score (equal-likelihood) contour about $\hat{\beta}$ for which the sum first equals or exceeds $1 - \alpha$.

The C_{R_p}'s are approximated by counting the number of combinations giving the same score encountered during the grid search of $\theta\phi\lambda$. The C_{R_p}'s and the corresponding likelihoods are stored in order of descending magnitude from R_{max}. Thus, K is the sum of this entire list. However, in practice, the summation is stopped when additional terms cease to contribute significantly. It is assumed that terms for $R_p < .5N$ never contribute significantly. The R_p on the boundary of F is found by summing this list again until the sum first equals or exceeds .95K.

Interpreted as a classical Bayesian confidence region, F is based on an a priori $Pr(\bar{\beta})$ which is constant and equal for each of the events $(\bar{\beta} \epsilon A_j)$. That is, the a priori probability associated with each CCR is the same; the occurrence of $\bar{\beta}$ in any CCR is a priori as probable as the occurrence of $\bar{\beta}$ in any other CCR. Interpreted as a percent-likelihood region or a region based on fiducial probability, no a priori probability need be brought in.

Another possibility which suggests itself is that the a priori $\phi(\bar{\beta})$ be uniform over the volume of the parameter space so that the a priori probability associated with a CCR is proportional to its volume. Then the sums for K and F have the form $\sum_\ell P^{R_p} (1-P)^{N-R_p} d\Omega_\ell$ where the sum is over the $\theta\phi\lambda$-grid points ℓ and where $d\Omega_\ell$, the differential element of volume in the $\phi\theta\lambda$-coordinates, is proportional to $\frac{1}{8} \sin\theta_\ell \, d\theta_\ell d\phi_\ell d\lambda_\ell$. This procedure is identical to the conventional treatment when all variables are continuous.

The mixture of discrete and continuous variables offers the possibility of discretization of the parameter space. Although

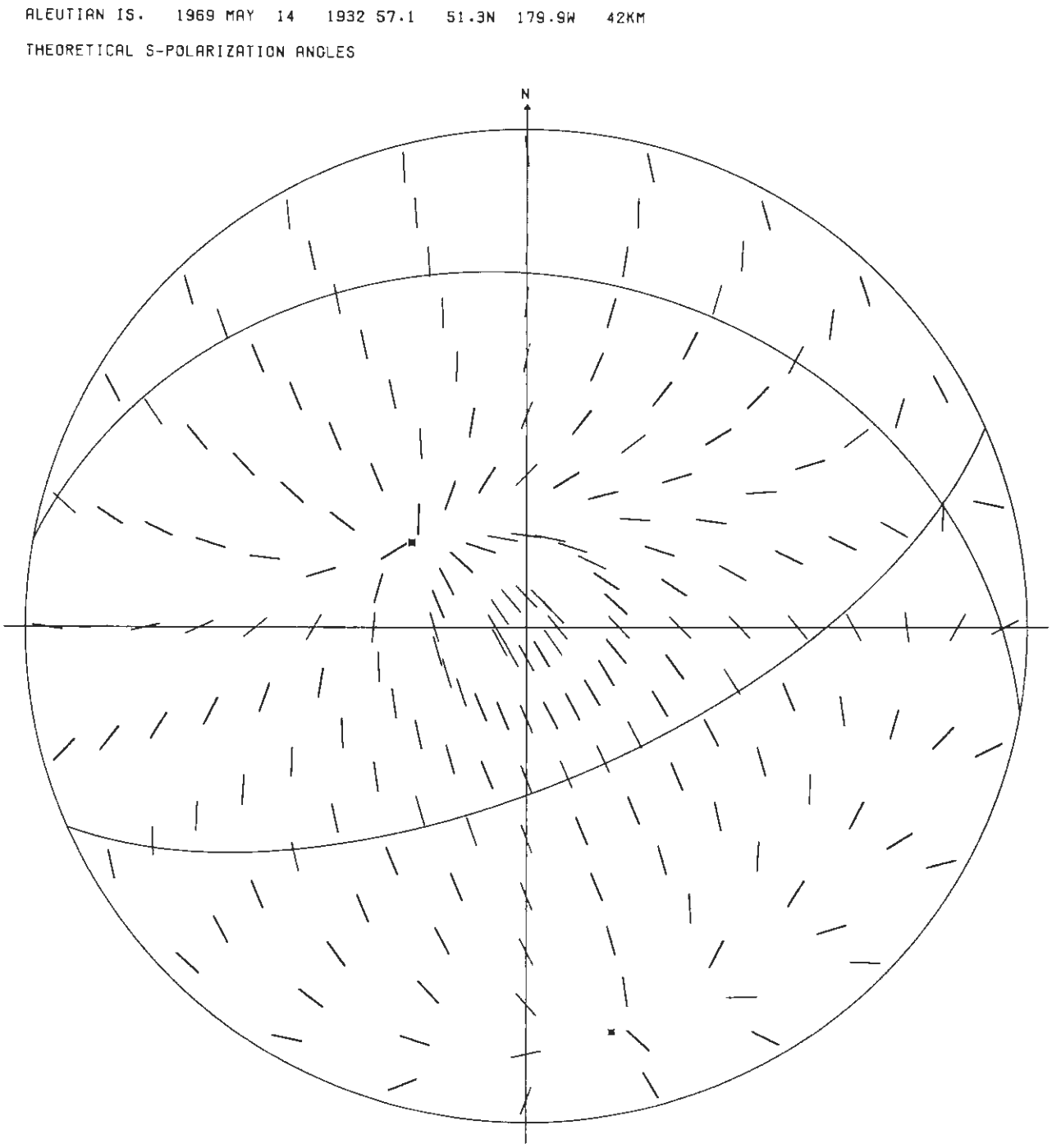

Figure 7.--Computer plot of the predicted S-polarization angles for the combined P- and S-wave solution.

alternative choices of an a priori distribution are more or less expected in the construction of classical Bayesian confidence regions, the same alternative--discretized versus continuous parameter space--is present in the percent-likelihood and fiducial probability interpretations as a different way of defining the likelihood summation they involve.

The summation over a discretized parameter space is closer to the summations that would be involved in the extension of the non-Bayesian strict confidence regions to the mixed case. An intuitively attractive feature of this procedure is that F-regions determined in this manner never let volume override score. Thus, even small volumes that have good scores are detected, while an a priori probability based on area could submerge these valid alternative solutions in favor of those having large constant-score volumes. Because the specific volumes are an accident of the distribution of stations and because small volumes reflect higher definition, it seems entirely unreasonable to reject these isolated alternatives on the basis of the subjective opinion, unsupported by data, that small areas are less likely to contain the solution than are large areas.

The guiding philosophy of the minimal interpretation of a percent-likelihood region is adopted here. No probability judgments not implicit in the data and their likelihood are introduced. The data should be "allowed to speak," unimpeded by the imposition of unnecessary probability judgments not supported by data. In fact, experience so far has shown that small alternative maximum-(or near maximum-) score regions frequently occur and are geophysically reasonable. The discretized parameter-space summation is compatible with exact regions and makes possible exact fiducial regions because the summations involve counting of combinations only rather than integrals of volume over regions with complicated boundaries.

An older procedure, since discarded, assumed that all $\binom{N}{R}$ possible combinations of stations producing a given score R could, in fact, be realized by variation of $\overline{\beta}$. This reduced the summations involved in F to cumulative binomial distribution functions which could be approximated for large N by the cumulative normal.

Because of the discrete grid sampling, the present count of C_{R_p} is approximate. However, only the rate of variation of C_{R_p} as a function of R_p is important because any constant factor cancels out in the "normalization" through division by K. We have found this approximate procedure, implemented with the aid of some programming finesse, to be a practical and satisfactory method which produces reasonable fiducial regions. Future efforts may involve further evolution of the method to provide exact regions, more efficient programs, and consistent procedures for fiducial regions on combined P-wave and S-wave solutions.

Unlike many routine adjustment problems, the computational details of exact small-sample fiducial regions for P-wave (and S-wave) first motion data are surprisingly sensitive to subtle (not to say obscure) variations in the statistical argument. This fact has motivated us to present the above discussion of general approaches to fiducial regions to provide an appreciation of the subtleties involved and a basis for further discussion.

V. DATA PROCESSING

The preceding sections of this report discuss some statistical techniques for aiding the evaluation of a focal plane solution and procedures for finding that solution from "observed" data projected on the focal sphere. How these observed data are corrected and projected to the focal sphere has much to do with the accuracy of the solution and with the meaningfulness of the statistics applied. All procedures currently used to get observed data on the surface of the focal sphere are fairly straightforward. However, if the accuracy of the solutions is to be improved, then at least some of the techniques for analyzing the observed data will surely have to be improved.

The actual process of generating a solution is summarized in the flow diagram in figure 8. The first step is to gather the seismograms to be used. These are usually obtained as 70-mm film chip copies from the National Geophysical Data Center in Asheville, N.C. For large, important earthquakes, a request is sent to cooperating seismograph stations throughout the world for the loan of their original seismograms of the earthquake.

The direction of first motion of the P-wave is read from the seismogram. When available, both long-period and short-period records are read. The first motions taken from long-period records are usually used in the solution; however, sometimes stations having only short-period seismographs may be located in critical positions relative to the mechanism. For this reason, it is sometimes desirable to supplement long-period data with short-period observations. The relative merit of using long- and short-period P-data is under study; this is one reason for reading both.

The stations having long-period records which are in the proper distance range are studied to see if they have usable S-wave data. If so, horizontal components of long-period seismograms from these stations are digitized in the vicinity of the S-wave arrival, using an x,y digitizer which records the output directly on magnetic tape. The digitizing begins at the start of a time mark and ends at another time mark. In addition to the x, y values, a code number is recorded. This code number is used to indicate the component, the beginning of a time mark, and the start of the S-wave. The beginning of the S-phase is picked

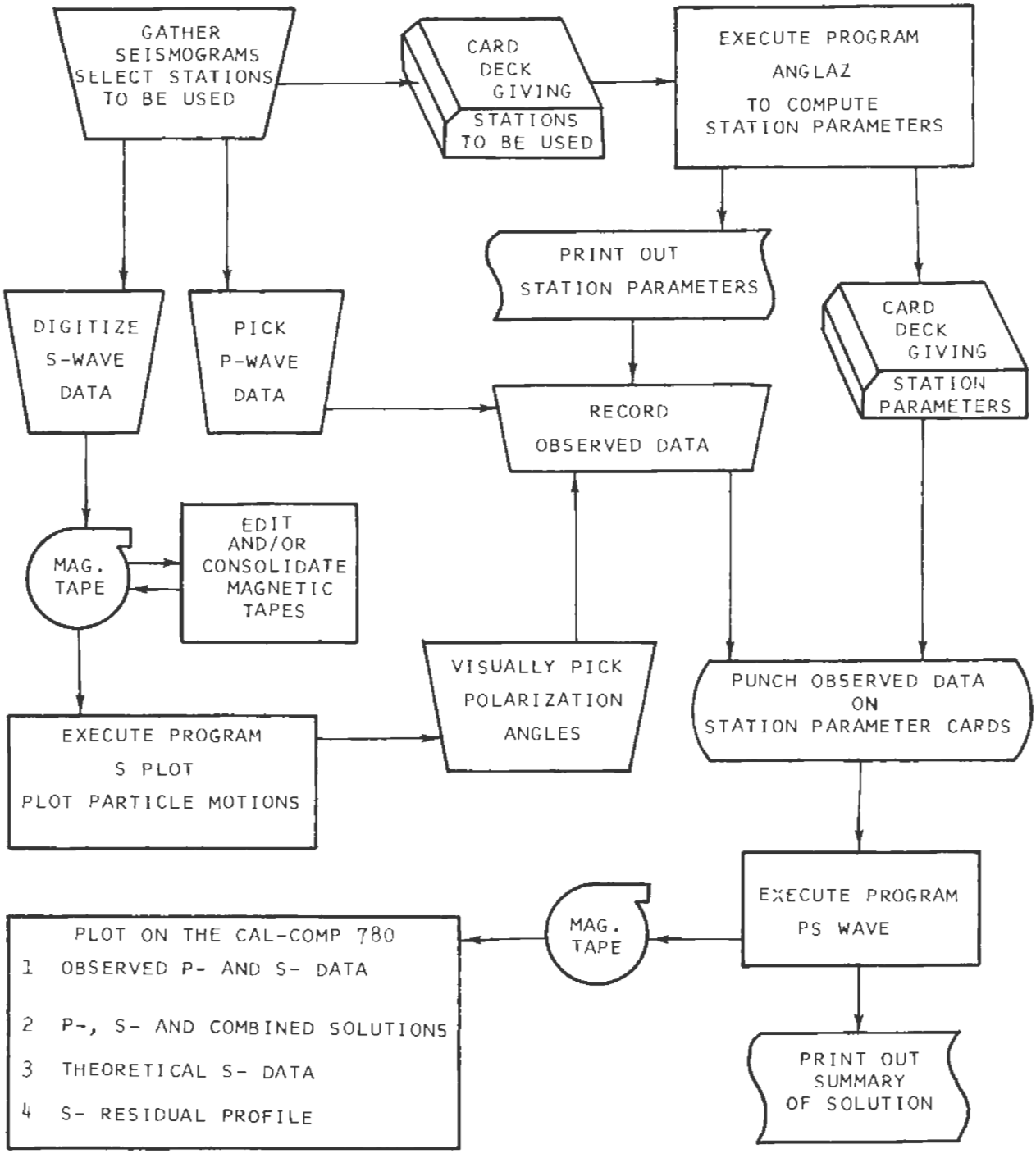

Figure 8.--System flow chart showing the operations and computer programs involved in computing focal mechanisms.

by the operator digitizing the record. When desirable, an operational computer program is available to edit or consolidate these tapes. Another program, SPLOT, will read these tapes and plot the digitized seismograms on a Cal-Comp plotter. These plots give back the data in the form of seismograms at the scale at which they were digitized for visual comparison with the original records. In addition, on the same execution run, a plot of the horizontal-particle motion for a 30-second portion of the record is given, beginning slightly before the start of the S-wave that was picked by the operator. The S-polarization angle is selected from this plot and recorded.

The observed polarization angle to be used as input to the program for computation of the focal mechanism solution is then computed from the relation:

(S-polarization angle) = (back azimuth) - (S-orientation).

It might be noted that while the S-orientation is measured positive north through east in the distance range normally used, the S-polarization angle is positive in the opposite sense because of the curvature of the ray path. At current writing, only stations at distances of between 40° and 80° are used, and no crustal corrections or surface interactions are computed for the polarization angles. This should not produce an error of more than 6° in the angle (Nuttli and Whitmore 1962). More experience with a larger number of solutions may help to determine if corrections are necessary.

Some typical S-wave particle-motion plots have been shown by Nuttli and Whitmore; because many plots of the authors are similar to theirs, we have not shown any here. However, we have noted dertain patterns which occur repeatedly. Standard deviations of 15° to 20° have not been unusual among the solutions we have run, and the same is true of those presented by Stauder (June 1968). Surface corrections in the distance range used can only account for a maximum error of 6°, so improvement would seem to be possible in the analysis of S-wave motion and the selection of S-polarization angles.

Coordinates $(\bar{x}_i, \bar{y}_i, \bar{z}_i)$ of an observation on the focal sphere have been obtained from the geographic coordinates of the station and hypocenter by projecting the observation back along the ray path. Those stations which are to be used in the solution are keypunched onto IBM cards. These cards are then used as input to the computer program ANGLAZ, shown in figure 8, which computes the necessary station parameters needed in the focal plane solution. The program ANGLAZ was developed to make these calculations using subroutines provided by Dr. E. P. Arnold. For each station, this program computes the distance, azimuth and back azimuth, predicted arrival time, and angle of departure of the ray at the source for both P- and S-waves. The angle of departure

ALEUTIAN IS. 1969 MAY 14 1932 57.1 51.3N 179.9W 42KM

P WAVE SOLUTION

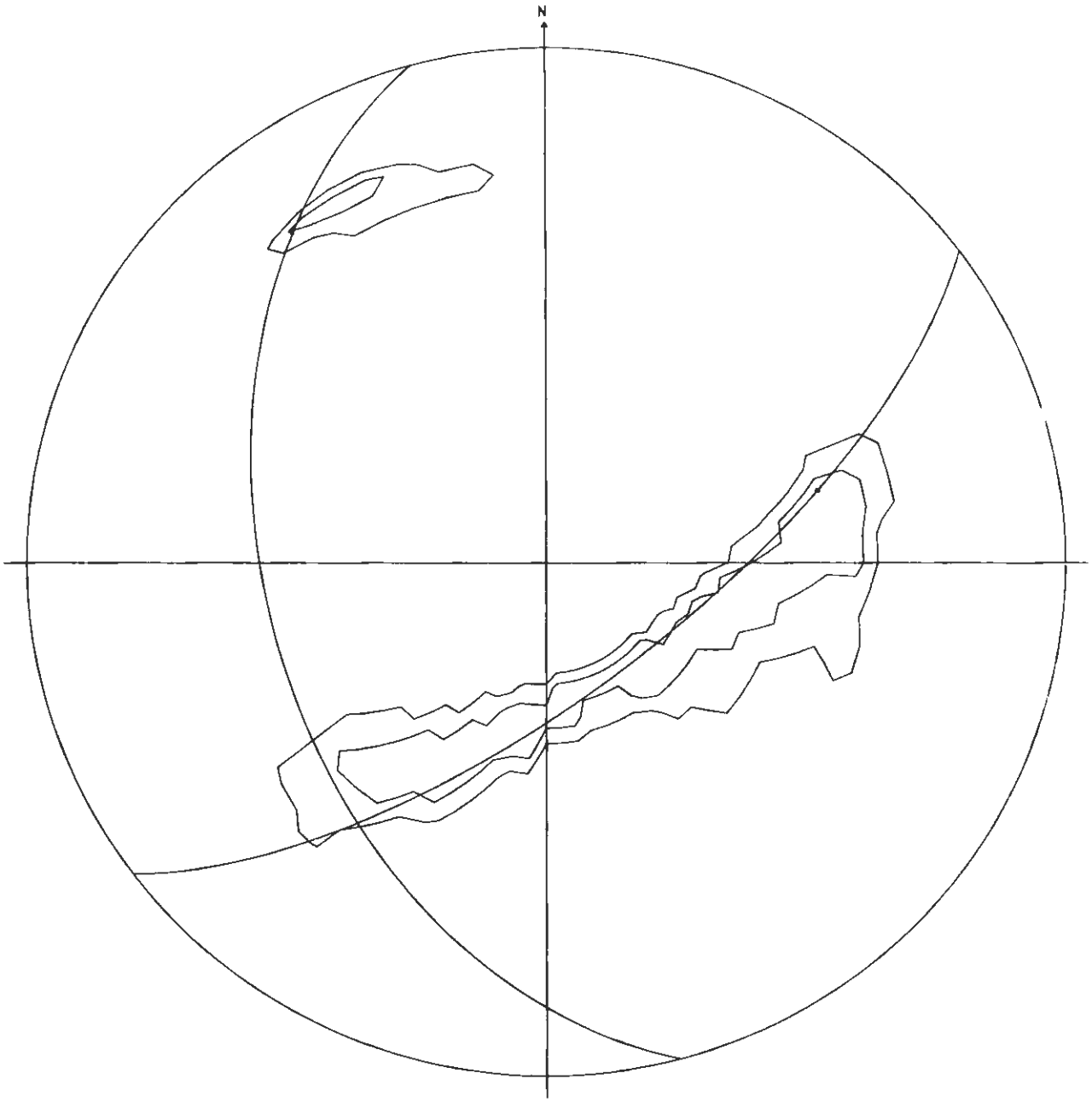

Figure 9.--Computer plot showing the 96.5-percent fiducial limit for the P-wave solution.

is computed using the JB (Jeffreys-Bullen) travel-time tables. Nine points are interpolated from the tables, using Gregory's formula with second forward differences (Jeffreys and Jeffreys 1956) to get the travel-time at the appropriate depth. Gregory's formula differentiated with second forward differences is then used on these points to compute the ray parameter $p = \frac{\partial T}{\partial \Delta}$. The angle of departure i_o is then computed in the normal closed form $i_o = \arcsin(p\frac{v}{r})$ where p is the ray parameter, v is the velocity at the hypocenter, r is the radius to the hypocenter, and all are converted to appropriate units. For stations within an arc-distance of about 1.5°, another program provided by Dr. E. R. Engdahl is used to make these same computations for P-waves because the travel-time curves used in this system are more finely divided for the crust near the epicenter. Both programs use JB travel-times and velocities. These programs print the parameters and also punch cards with the same data. The observed P-wave first motions and S-polarization angles are then keypunched on these parameter cards and are used as input to the computer program PSWAVE.

The actual calculation of the solution and the contouring of the fiducial limits are done by the program PSWAVE (see figure 8) which is operational on a CDC-6600 computer. This program will find either a P-wave, an S-wave, or a combined P- and S-wave solution.

When program PSWAVE has determined the best solution and the statistical constants, it contours the fiducial limits and plots the data as shown in figures 5 and 6. The program actually writes a magnetic tape containing the proper instructions for a Cal-Comp plotter to draw the diagrams. Figures 5 through 7 and 9 through 12 were all computer-generated by the program PSWAVE. A summary of the solution is printed out in table 1. Because of the large number of increments needed for precise solutions, some of the program has been done in assembly language to increase the speed of the calculations. Moreover, because of the length of the programs, they are not printed here.

The Aleutian earthquake of May 14, 1969, was selected as a good example of the procedures presented here. The P-wave data plotted in figure 5 were all selected from long-period instruments of the worldwide system. Short-period records were also selected and agreed in general with the long-period data. At least there is no acceptable short-period solution which is significantly different from the long-period results. The evaluation of this solution was carried out on a five-degree grid. The solutions selected by the program PSWAVE are shown in table 1. The orientation of the mechanism is specified by the x- and y-axes of the focal mechanism coordinate system, and the contours showing the fiducial regions are limits on the allowable variation of these axes. Therefore, it is convenient to specify the solution by giving the strike and dip of these vectors. The strikes

Table 1.--Earthquake focal plane solution.

Aleutian Is. 1969 May 14 GMT 19 hrs. 32 mins. 57.1 secs.
51.3N. 179.9W. 42 km.

	Pole of A-Plane		Pole of B-Plane	
	Strike	Dip	Strike	Dip
P-wave solution	75.00	45.00	322.79	69.30
S-wave solution	180.00	30.00	337.20	61.98
Combined solution	190.00	30.00	336.31	64.36

.....

P-Wave Solution

Total observations	62
Agreeing observations	62
Disagreeing observations	0
Computed probability $\hat{P}$	1.0
Fiducial limit (based on P = .95)	96.5%
Fiducial limit R_p	61

S-Wave Solution

Total observations	19
Sum of squared residuals R_s	5,002.5
Standard deviation	17.68
Fiducial limit	95.0%
Fiducial limit R_s	8,041.5

Combined Solution

Agreeding P-wave first motions R_p	61
S-wave sum of squared residuals R_s	6,260.9

Weighting constants (a = 2.944, b = .00160)

ALEUTIAN IS. 1969 MAY 14 1932 S7.1 51.3N 179.9W 42KM

S WAVE SOLUTION

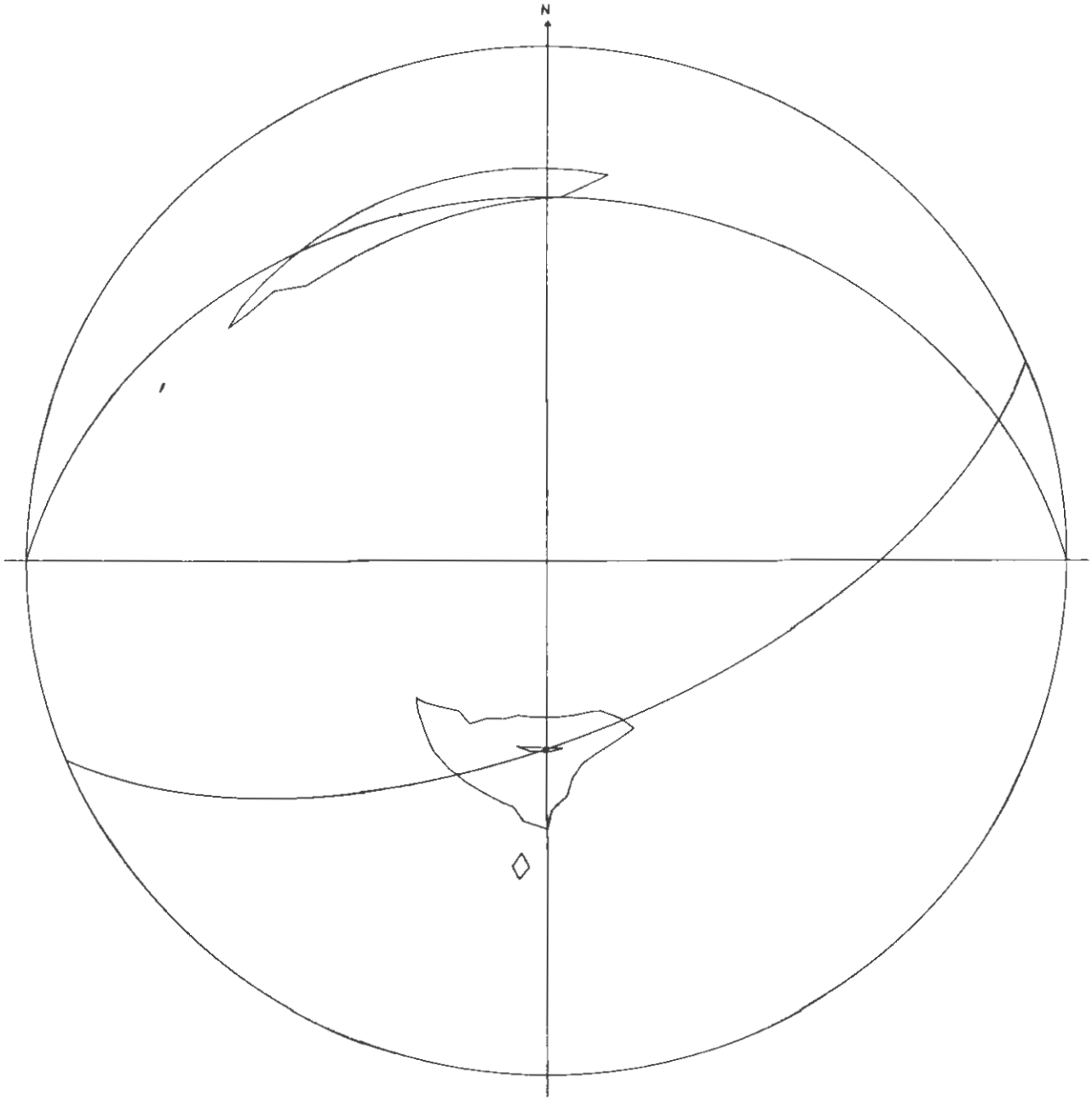

Figure 10.--Computer plot of the 95-percent fiducial limits for the S-wave solution.

and dips given in the computer plots and in table 1 all refer to vectors normal to the planes, that is, the poles of the focal planes. When the solution is not unique, only the first one found is printed in the summary table, but the others can be seen from the contoured regions, as shown in figure 9 for the P-wave solution. Because the S-wave data produce a continuous error surface, it is less likely that there will be solutions of equal R_s.

To aid in seeing the trend of the S-wave error surface, in addition to the 95-percent fiducial limit, an arbitrary contour is plotted close to the S-wave minimum as can be seen in figure 10. An additional plot giving a profile of the sum of squares of S-wave residuals is shown in figure 12. With the first plane held fixed at the value given by the combined solution, the second plane is rotated through 180° to generate figure 12, showing how well defined the minimum is for the S-wave data.

Statistics which will give the value of a confidence limit for the combined solution have not been presented here. However, it is still desirable to see the trend of the likelihood surface in the vicinity of the solution. To display the trend of the likelihood surface, an arbitrary interval is selected and equal-score curves are contoured at equal intervals from the maximum. These are shown in figure 11 for the May 14 data.

The resulting solutions for the May 14 earthquake are quite good for illustrating the desirable properties of focal plane solutions. The fiducial limits for the individual P-wave and S-wave solutions overlap, and the combined solution falls within the limits of both. As can be seen in figure 9, the P-wave solution determines one pole very well while the second pole and its corresponding plane are free to move over a broad range of values. The S-wave fiducial limits in figure 10 have a much different pattern and restrict the limits of the second pole considerably.

The P-wave solution has no discriminating power within a constant-score region and the S-wave data are in good agreement with a portion of the P-wave fiducial region. Therefore, the combined solution seen in figure 11 is not much different from the S-wave solution shown in figure 10. So in this case, the S-wave data strongly influenced the selection of the final combined solution. If it were desired to increase the influence of the P-wave first motions, the assigned value of P could be increased to a value more nearly equal to 1.0 (P was assigned a value of 0.95 for this combined solution, as shown in table 1). The final combined solution has allowed one of the P-wave observations to be considered in error, giving 98.5-percent agreement for the P-wave data.

Thus, the combined solution represents a good solution to both the P- and S-wave data, and the two are in good agreement. It is interesting to note that while the S-wave data have had a strong influence in finding the final solution for the May 14

ALEUTIAN IS. 1969 MAY 14 1932 57.1 51.3N 179.9W 42KM

COMBINED SOLUTION

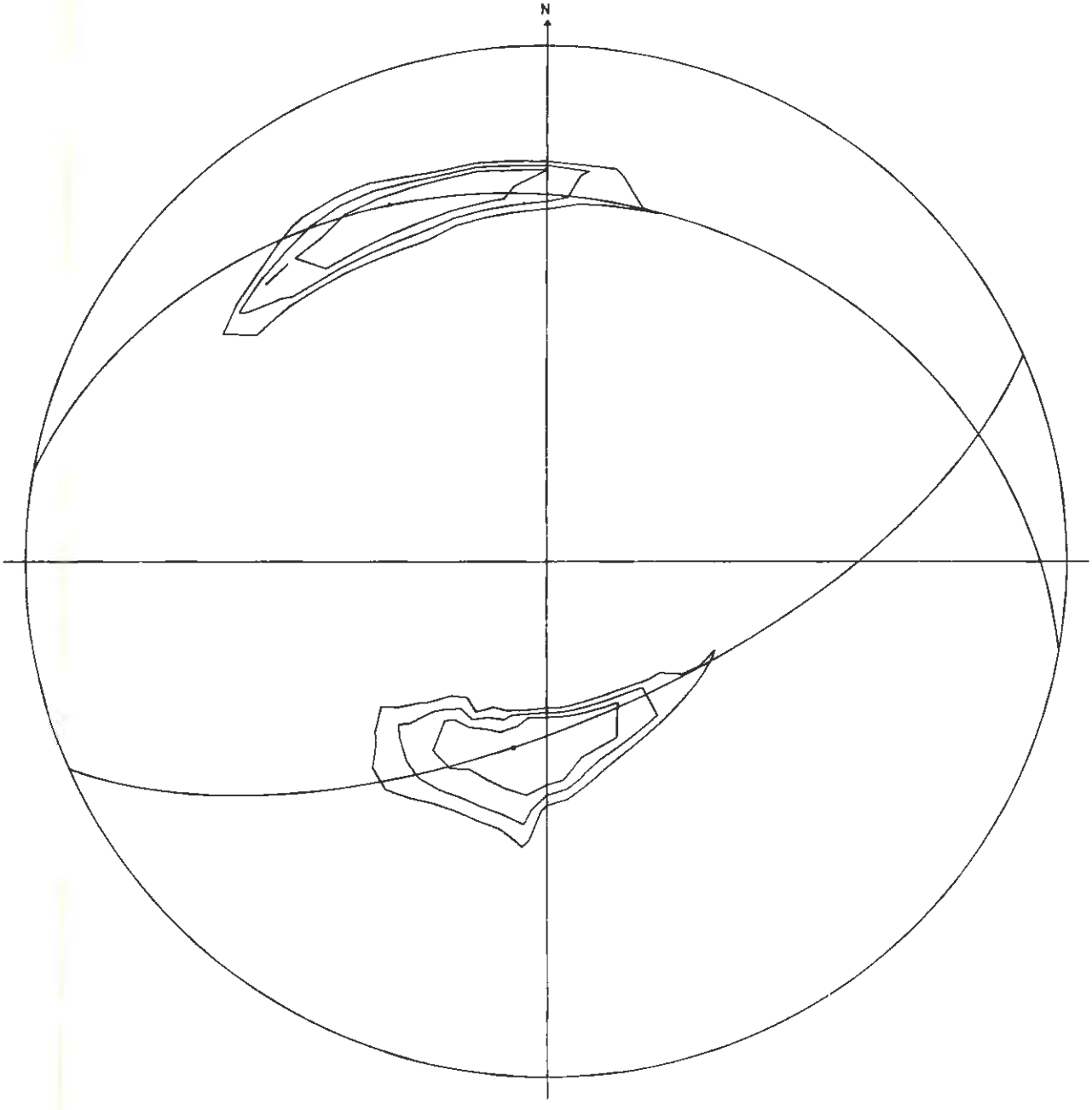

Figure 11.--Computer plot showing the best combined P- and S-wave solution and contours of equal likelihood around the maximum-likelihood solution.

earthquake, it does not always happen that way. The authors have seen cases where, because of the poor distribution of the S-wave data and the better distribution of the P-wave data, the S-wave polarization angles have had very little effect on the solution. This circumstance shows the value of using both P- and S-wave data with proper weighting as developed above.

VI. CONCLUSION

The system described here provides NOAA with the capability needed for routine computer calculation of focal mechanisms based on P-wave data, S-wave data, and combinations of the two. The presentation of fiducial limits makes possible the evaluation of the quality and allowable variation of any P-wave or S-wave solution. As these limits have the same meaning from one solution to another and are presented on an equal-area grid, the relative quality of solutions is easy to evaluate.

The P-wave, S-wave, and combined P- and S-wave solutions are all based on straightforward statistical arguments using the method of maximum-likelihood estimation. More complex statistical models have been rejected in favor of a direct approach based on clearly stated simplifying assumptions. Our experience with unpublished solutions has convinced us that this direct approach is adequate and produces results that are reasonable and intuitively satisfying.

The new graphical-geometric method also presented here for finding and constructing the exact boundaries of P-wave solutions possesses significant advantages over conventional P-wave solutions. This method has already proved to be a valuable supplement to the numerical grid construction of boundaries; it is also useful for checking programs and for studying particular earthquakes in detail. New computer programs based on this method are planned; these are expected to provide exact limits for the P-wave solution and for all significant alternatives more efficiently than the present program using a grid search.

A study of the May 14, 1969, Aleutian earthquake shows clearly the desirable properties of solutions constructed by these aforementioned procedures. In particular, the agreement found between the P-wave first motions and the S-polarization angles supports the concept of combined solutions.

The emphasis in this report has been on the methodology of focal plane determinations and on the derivation and implementation of solution methods. Currently operational programs constitute a tool for the extension of work in: (1) routine calculation and cataloging of mechanisms; (2) local and regional

ALEUTIAN IS. 1969 MAY 14 1932 57.1 51.3N 179.9W 42KM

INITIAL PLANE STRIKE 190.00 DIP 30.00

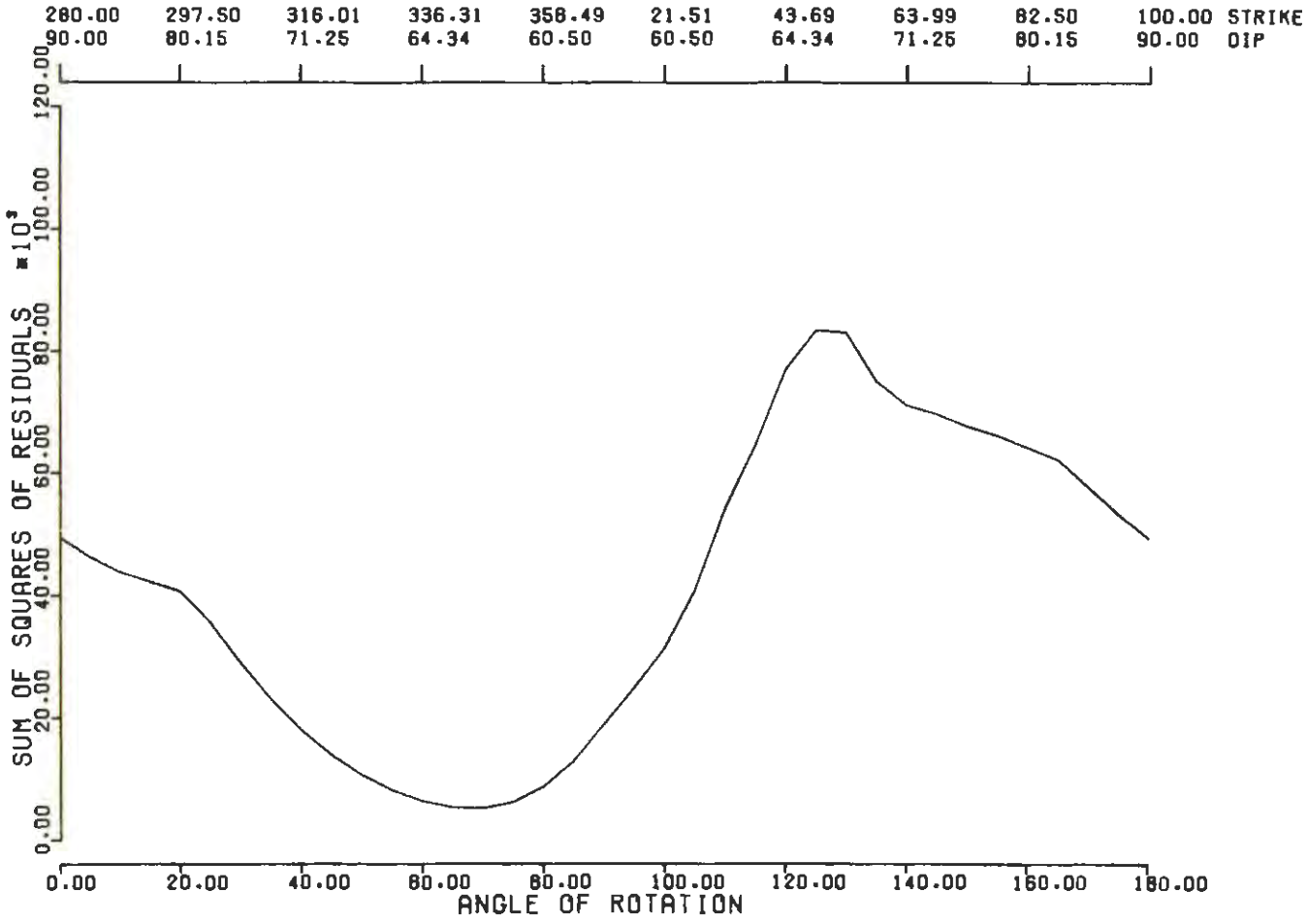

Figure 12.--Computer plot showing a profile of the S-wave score with the initial plane held fixed at the value given.

seismicity studies, including composite solutions; and (3) detailed studies of important earthquakes. The technique of presenting contoured fiducial limits on a projection of the lower half of the focal sphere makes possible an evaluation of the geophysical significance of all alternate solutions rather than an evaluation relying upon a single solution by a particular investigator. Conventional solutions without fiducial limits permit variation only if the data are reanalyzed, that is, if the fault plane solution is repeated in a search for alternatives consistent with the data. Because fiducial regions show all possible alternatives, a particular investigator's choice of the focal planes, which may have been influenced by his conception of the expected geophysical situation, can no longer exclude other choices by subsequent investigators.

The methods described herein can be extended in several directions which constitute the goals of current development work. These goals include efficient computation of combined P- and S-fiducial regions, quality factors for "one number" ranking of quality, and programs to exploit the above described construction for exact P-wave contours and solution. A parallel and continuing effort is devoted to the search for improvements in S-polarization angle determinations because good observational data are essential for good results.

ACKNOWLEDGMENTS

The development of this computer system has been accomplished with the aid of many discussions. The authors wish to thank Dr. E. P. Arnold of the International Seismological Centre for many helpful discussions and for providing operational computer subroutines and programming assistance. Dr. E. R. Engdahl of the Seismology Division of NOS has also kindly provided computer support and operational programs over a long period of time.

REFERENCES

Benjamin, J. R., and Cornell, C. A., Probability, Statistics, and Decision for Civil Engineers, McGraw-Hill Book Co., Inc., New York, 1970, 684 pp.

Cramér, H., Mathematical Methods of Statistics, Princeton University Press, Princeton, N. J., 1946, pp. 498-520.

Draper, N. R., and Smith, H., Applied Regression Analysis, John Wiley and Sons, Inc., New York, 1966, 407 pp.

Graybill, F. A., An Introduction to Linear Statistical Models, Vol. 1, McGraw-Hill Book Co., Inc., New York, 1961, 463 pp.

Hamilton, W. C., Statistics in Physical Science: Estimation, Hypothesis Testing, and Least Squares, The Ronald Press Co., New York, 1964, pp. 124-145.

Hirasawa, T., "A Least Squares Method for the Focal Mechanism Determination from S-Wave Data; Part I," Bulletin of the Earthquake Research Institute, Vol. 44, 1966, pp. 901-918.

Isacks, B., Oliver, J., and Sykes, L. R., "Seismology and the New Global Tectonics," Journal of Geophysical Research, Vol. 73, No. 18, September 15, 1968, pp. 5855-5899.

Jeffreys, H., Sir, and Jeffreys, B. S., Methods of Mathematical Physics, 3d. ed., Cambridge University Press, Cambridge, England, 1956, pp. 265-266.

Kasahara, K., "Computer Program for a Fault-Plane Solution," Bulletin of the Seismological Society of America, Vol. 53, No. 1, January 1963, pp. 1-13.

Knopoff, L., "Analytical Calculation of the Fault-Plane Problem," in "A Symposium on Earthquake Mechanism," Publications of the Dominion Observatory, Vol. 24, No. 10, 1961, pp. 309-315.

Mood, A. M., and Graybill, F. A., Introduction to the Theory of Statistics, 2d. ed., McGraw-Hill Book Co., Inc., New York, 1963, 443 pp.

Nuttli, O., and Whitmore, J. D., "On the Determination of the Polarization Angle of the S Wave," Bulletin of the Seismological Society of America, Vol. 52, No. 1, January 1962, pp. 95-107.

Rao, C. R., Linear Statistical Inference and Its Applications, John Wiley and Son, Inc., New York, 1965, 522 pp.

Stauder, W., "S Waves and Focal Mechanisms: The State of the Question," Bulletin of the Seismological Society of America, Vol. 50, No. 3, July 1960, pp. 333-346.

Stauder, W., "Mechanism of the Rat Island Earthquake Sequence of February 4, 1965, with Relation to Island Arcs and Sea-Floor Spreading," Journal of Geophysical Research, Vol. 73, No. 12, June 15, 1968, pp. 3847-3858.

Stauder, W., "Tensional Character of Earthquake Foci Beneath the Aleutian Trench with Relation to Sea-Floor Spreading," Journal of Geophysical Research, Vol. 73, No. 24, December 15, 1968, pp. 7693-7701.

Stauder, W., and Bollinger, G. A., "The S Wave Project for Focal Mechanism Studies, Earthquakes of 1962," Bulletin of the Seismological Society of America, Vol. 54, No. 6, Part B, December 1964, pp. 2199-2208.

Stevens, A. E., "S-Wave Earthquake Mechanisms Equations," Bulletin of the Seismological Society of America, Vol. 57, No. 1, February 1967, pp. 99-112.

Stevens, A. E., and Hodgson, J. H., "A Study of P Nodal Solutions (1922-1962) in the Wickens-Hodgson Catalogue," Bulletin of the Seismological Society of America, Vol. 58, No. 3, June 1968, pp. 1071-1082.

Sykes, L. R., "Mechanism of Earthquakes and Nature of Faulting on the Mid-Oceanic Ridges," Journal of Geophysical Research, Vol. 72, No. 8, April 15, 1967, pp. 2131-2153.

Udias, A., and Baumann, D., "A Computer Program for Focal Mechanism Determination Combining P and S Wave Data," Bulletin of the Seismological Society of America, Vol. 59, No. 2, April 1969, pp. 503-519.

Udias, A., "A Least Squares Method for Earthquake Mechanism Determination Using S-Wave Data," Bulletin of the Seismological Society of America, Vol. 54, No. 6, December 1964, pp. 2037-2047.

Wickens, A. J., and Hodgson, J. H., "Computer Re-evaluation of Earthquake Mechanism Solutions 1922-1962," Publications of the Dominion Observatory, Vol. 33, No. 1, 1967, 560 pp.

Wilks, S. S., Mathematical Statistics, John Wiley and Son, Inc., New York, 1962, 644 pp.

Zhelankina, T. S., Keilis-Borok, V. I., Pisarenko, V. F., Piatetskiĭ-Shapiro, I. I., "Opredelenie Mekhanizma Zemletriasenii Na Tsifrovoĭ Elektronnoĭ Vychislitel'Noĭ Mashine" (Determination of Earthquake Mechanism on Electronic Computers), in Russian, in press.

LIST OF NOTATIONS

$\overline{\chi}$ vector of observables

$\overline{\beta}$ vector parameters

$\hat{\beta}_{ML}$ maximum likelihood estimate of parameters

$\phi(\overline{\chi})$ probability density function (p.d.f.) of $\overline{\chi}$

$\phi(\overline{\beta})$ p.d.f. of $\overline{\beta}$ (an "a priori" p.d.f.)

$\phi(\overline{\chi}|\overline{\beta})$ p.d.f. of $\overline{\chi}$ conditional upon $\overline{\beta}$

$\phi(\overline{\chi},\overline{\beta})$ joint p.d.f. of $\overline{\chi}$ and $\overline{\beta}$

$\phi(\overline{\chi};\overline{\beta})$ p.d.f. of $\overline{\chi}$ functionally dependent upon parameters $\overline{\beta}$

$Pr(\overline{\chi}_p|\overline{\beta})$ probability function of $\overline{\chi}_p$ conditional upon $\overline{\beta}$

$\phi r(\overline{\chi}_p,\overline{\chi}_s|\overline{\beta})$ mixed probability function of $\overline{\chi}_p$ (discrete variables) and $\overline{\chi}_s$ (continuous variables)

$L(\overline{\beta};\overline{\chi})$ likelihood of $\overline{\beta}$

$L(\overline{\beta};\overline{\chi}_p,\overline{\chi}_s)$ likelihood of $\overline{\beta}$, based on $\phi r(\overline{\chi}_p,\overline{\chi}_s|\overline{\beta})$

$\overline{V}$ vector of residuals, v_i

v_i residual = difference between the value of an observable computed by a model and its observed value.

R_s "S-score" = $\overline{V}^t W \overline{V}$ (= $\sum_i v_i^2$ when $W = I$)

R_p "P-score" = number of agreements between predicted and observed senses of first motion of P-wave

R_{ps} "combined P-S score" = $aR_s + bR_p$

$$\left.\begin{aligned} a &= \ln\left(\frac{P}{1-P}\right) \\ b &= \frac{1}{2\sigma^2} \end{aligned}\right\} \text{coefficients in } R_{ps}$$

$\overline{R}_p = N - R_p$ = number of disagreements

P probability of reading a P-wave first motion correctly

σ standard error of a single polarization angle of weight one.

c a constant in the multivariate normal p.d.f. = $(\det W)^{1/2}$ $(2\pi\sigma^2)^{-n/2}$

k a constant in the combined P-S likelihood function = cP^n

M covariance matrix $M = \{\sigma_{ij}\}$ where $\sigma_{ij} = \begin{cases} \sigma_i^2 & \text{variance } i = j \\ \sigma_{ij} & \text{covariance } i \neq j \end{cases}$

W weight matrix = $M^{-1}\sigma^2$

I Fisher's information matrix = $- E\left\{\frac{\partial^2 \ln L}{\partial\overline{\beta}\partial\overline{\beta}}\right\}$

$A(\overline{\beta})$ acceptance region

$C(\hat{\beta})$ strict confidence region

K a constant appearing in the renormalized likelihood = $\sum_{\text{all}\,\overline{\beta}} L$

A a region in the parameter space

ε as in $\overline{\beta}\varepsilon A$ or $\overline{\beta}\not\varepsilon A$, read "is an element of"

ε polarization angle

ϕ azimuth clockwise from north

θ spherical colatitude

λ "third angle" parameter of rotation of B-pole about A-pole

$d\Omega_\ell$ element of volume in the parameter space

SCC acronym for "station complementary circle"

CCR acronym for "constant combination region"

(Continued from inside front cover)

C&GS 18. Submarine Physiography of the U.S. Continental Margins. G. F. Jordan, March 1962. Price $0.20.

*C&GS 19. Analytic Absolute Orientation in Photogrammetry. G. C. Tewinkel, March 1962.

*C&GS 20. The Earth as Viewed from a Satellite. Erwin Schmid, April 1962.

*C&GS 21. Analytic Aerotriangulation. W. D. Harris, G. C. Tewinkel, C. A. Whitten, July 1962 (Corrected July 1963)

*C&GS 22. Tidal Current Surveys by Photogrammetric Methods. Morton Keller, October 1963.

*C&GS 23. Aerotriangulation Strip Adjustment. M. Keller and G. C. Tewinkel, August 1964.

*C&GS 24. Satellite Triangulation in the Coast and Geodetic Survey, February 1965.

C&GS 25. Aerotriangulation: Image Coordinate Refinement. M. Keller and G. C. Tewinkel, March 1965. Price $0.15.

C&GS 26. Instrumented Telemetering Deep Sea Buoys. H. W. Straub, J. M. Arthaber, A. L. Copeland, and D. T. Theodore, June 1965. Price $0.25.

*C&GS 27. Survey of the Boundary Between Arizona and California. Lansing G. Simmons, August 1965.

*C&GS 28. Marine Geology of the Northeastern Gulf of Maine. R. J. Malloy and R. N. Harbison, February 1966.

C&GS 29. Three-Photo Aerotriangulation. M. Keller and G. C. Tewinkel, February 1966. Price $0. 35.

C&GS 30. Cable Length Determinations for Deep-Sea Oceanographic Operations. Capt. Robert C. Darling, June 1966. Price $0.10.

C&GS 31. The Automatic Standard Magnetic Observatory. L. R. Alldredge and I. Saldukas, June 1966. Price $0.25.

ESSA Technical Reports--C&GS

C&GS 32. Space Resection in Photogrammetry. M. Keller and G. C. Tewinkel, September 1966. Price $0.15.

C&GS 33. The Tsunami of March 28, 1964, as Recorded at Tide Stations. M. G. Spaeth and S. C. Berkman, July 1967. Price $0.50.

C&GS 34. Aerotriangulation: Transformation of Surveying and Mapping Coordinate Systems. Lt. Cdr. Melvin J. Umbach, August 1967. Price $0.25.

C&GS 35. Block Analytic Aerotriangulation. M. Keller and G. C. Tewinkel, November 1967. Price $0.55.

C&GS 36. Geodetic and Grid Angles--State Coordinate Systems. Lansing G. Simmons, January 1968. Price $0.10.

C&GS 37. Precise Echo Sounding in Deep Water. G. A. Maul, January 1969. Price $0.25.

C&GS 38. Grid Values of Total Magnetic Intensity IGRF--1965. E. B. Fabiano and N. W. Peddie, April 1969. Price $0.60.

C&GS 39. An Advantageous, Alternative Parameterization of Rotations for Analytical Photogrammetry. Allen J. Pope, September 1970. Price $0.30.

C&GS 40. A Comparison of Methods of Computing Gravitational Potential Derivatives. L. J. Gulick, September 1970. Price $0.40.

NOAA Technical Reports--NOS

NOS 41. A User's Guide to a Computer Program for Harmonic Analysis of Data at Tidal Frequencies. R. E. Dennis and E. E. Long, July 1971.

NOS 42. Computational Procedures for the Determination of a Simple Layer Model of the Geopotential From Doppler Observations. Bertold U. Witte, April 1971. Price $0.65.

NOS 43. Phase Correction for Sun-Reflecting Spherical Satellite. Erwin Schmid, July 1971.