BULLETIN OF THE EARTHQUAKE
RESEARCH INSTITUTE
Vol. 44 (1966), pp. 901-918

47. A Least Squares Method for the Focal Mechanism Determination from S Wave Data; Part I.*

By Tomowo HIRASAWA,

Geophysical Institute, Faculty of Science, The University of Tokyo.

(Read July 20, 1965.—Received June 29, 1966.)

Abstract

A least squares method for the focal mechanism determination is given by use of S wave data, where the double couple hypothesis is assumed.

The substantial idea of the method is based upon the assumption that a finite set of the residuals of the polarization angles of S waves for a particular earthquake is sampled from a population of the normal distribution, where the residual is defined as deviation of the observed polarization angle from that expected theoretically at an individual station. Because of the non-linearity of the problem the sum of squares of the residuals has secondary minima in some cases. For such cases the use of the statistical test for the significance of difference of the population variances is proposed in order to distinguish the most probable solution independent of the P wave observations.

1. Introduction

Up to 1960 the fault plane solution for an earthquake had been determined visually by use of the compression-rarefaction distribution of the initial motion of P waves in general. The polarization angles of S waves were used for selecting an acceptable source model for an earthquake from two hypothetical models, that is, a single couple of body forces (type I) and a double couple (type II).

It seems to be very difficult to find analytical criteria for the best fit required for the numerical determination of the focal mechanism from the P wave data, since there is no information available from observations, in general, except the sign of initial motion of the P waves. Although there are such difficulties, Knopoff (1961 *a*, *b*) has given an analytical

* Communicated by H. HONDA.

method to solve the problem. He introduces a kind of probability function defined so that its maximum represents the maximum likelihood solution. According to his definition, the probability that the reported sign is correct depends upon the amplitude of the P wave expected theoretically from the solution to be estimated.

Kasahara (1963) has improved the method. The amplitude rigorously calculated is used in his computer program, while Knopoff uses the approximate amplitude which depends upon the relative distance of the station from the two nodal planes of the P wave on a particular projection of the focal sphere. Furthermore, Kasahara introduces the weight assigned to each observation depending on the past reliability of station as reported by Hodgson and Adams (1958). He has applied his computer program to some earthquakes and obtained satisfactory results. It does not seem, however, that the probability function ingeniously proposed by Knopoff has been so satisfactorily justified that the standard error presented by him is statistically meaningful. Recently Hodgson and Wickens (1965) presented another numerical method for solving the problem.

On the contrary the problem concerning S wave data is rather simple since the continuous quantity of polarization angle of the S wave can be made use of. There seem to be, however, two difficulties in the case of S wave data. The first is that the two hypothetical source models (single couple and double couple) for the earthquake focus give different radiation patterns to the S wave. This difficulty, however, has almost been removed, and it seems that most of earthquakes, at least, can be explained by the double couple hypothesis (*cf.* Stauder and Bollinger; 1964). The other difficulty is that the analytical expression of the polarization angle derived from the theoretical model is very complicated especially for the double couple model, the complexity leading investigators to overlook what quantity should be minimized in applying the method of least squares to the problem.

As early as 1958 Adams presented a numerical method for the problem assuming the single couple hypothesis. He solved equations with respect to unknowns for all possible pairs of observations and averaged these solutions to determine the unknowns.

Stevens (1964) has given a least squares method for the double couple model. A family of non-linear equations with respect to three independent unknowns is solved by the following numerical method. A table is prepared of sets of values of normalized direction cosines to represent 3240 orientations of the double dipole axis. The computer evaluates the quanti-

ty $(1/N)\sum R_N^2$ for each set of axes, where R_N is residual for the Nth station and a given set of axes. The best solution is the one for which this mean sum of the squares of the residuals is a minimum. The computer selects from the table the 10 best orientations. For each of these, the computer again computes the above quantity for 45 possibilities in the neighborhood. If the 10 best of the final 450 possible solutions are closely alike in orientation, then the best solution is taken to be the orientation of the double dipole.

Although the method is certainly a kind of least squares method, it should be noted that the quantity which is minimized in her method is not the sum of squares of deviations of the observed polarization angles from the theoretical ones expected for the solution. The method is very advantageous, however, since it does not require any preliminary information about the solution of the earthquake concerned and gives a solution independently of the P wave data.

Udias (1964) has presented a numerical method for the double couple model. The method is based upon a preliminary estimate or graphical solution of one of the two P-nodal planes. The first least squares procedure in his method varies the position of the pre-estimated plane so as to give a minimum scatter of the poles of the second nodal plane, and a best position of the first nodal plane is determined. Now there is only one independent unknown to be estimated. Each observation can give a solution of the unknown parameter, and an average of the solutions thus obtained is the final solution. In many practical cases, one can determine one of the two P-nodal planes by use of the P wave data but not the other plane. In such cases the method may be very efficient for practical purposes.

Numerical methods for the single couple hypothesis were given both by Stevens (1964) and by Udias (1964). As is well known, for the single couple the forces at the focus are directed along a single axis, and the direction of the forces at the source and the S motion at a point on the focal sphere are co-planar. Consequent upon the co-planar property, the direction of the forces, or only one of the two P-nodal planes, can be determined from the polarization angle of the S waves defined irrespective of the sense of the S motion. Since there are only two independent unknowns, the problem of the focal mechanism determination is simpler than for the double couple. Moreover, the single couple is of little importance for the model of earthquake focus as was stated before. In the present paper, therefore, the method for the single couple model will not be given.

Stauder and Bollinger (1964) have presented the frequency distribution

of the residuals which are differences between the observed polarization angles and the theoretical ones expected from the double couple solution estimated graphically, superposing data obtained from thirty-six earthquakes. The frequency distribution suggests that it is a sample distribution from a population of the normal distribution. If a sample of the residuals defined above for a particular earthquake is assumed to be sampled from a population of the normal distribution, the maximum likelihood estimates of the unknowns are given by minimizing the sum of squares of the residuals of the observed polarization angles at the individual stations from the polarization angles calculated for the theoretical source.

The substantial idea of the method given in the present paper exists in the above assumption, and statistical discussions on the assumption will be made in the second paper.

2. Theory

A double couple source is formed by the superposition of two couples of equal forces with opposite moments. Let (x, y, z) be a right-handed coordinate system related to the orientation of the source mechanism. The Cartesian coordinates (x, y, z) are transformed into the spherical coordinates (r, θ, φ).

$$x=r\sin\theta\cos\varphi\,,$$
$$y=r\sin\theta\sin\varphi\,,$$
$$z=r\cos\theta\,.$$

When the double couple of body forces which are oriented along the y- and z-axes acts at the origin of the coordinates in an infinite elastic medium, the displacement components at a large distance from the source are written in the spherical coordinates as follows (*cf.* Honda; 1962);

$$\left.\begin{aligned}
\delta_{pr}&=\frac{K}{4\pi\rho}\frac{1}{v_p^3}\frac{u_r}{r}i\omega\exp i\omega\left(t-\frac{r}{v_p}\right);\quad u_r=\sin 2\theta\sin\varphi\,,\\
\delta_{s\theta}&=\frac{K}{4\pi\rho}\frac{1}{v_s^3}\frac{u_\theta}{r}i\omega\exp i\omega\left(t-\frac{r}{v_s}\right);\quad u_\theta=\cos 2\theta\sin\varphi\,,\\
\delta_{s\varphi}&=\frac{K}{4\pi\rho}\frac{1}{v_s^3}\frac{u_\varphi}{r}i\omega\exp i\omega\left(t-\frac{r}{v_s}\right);\quad u_\varphi=\cos\theta\cos\varphi\,,
\end{aligned}\right\}\qquad(1)$$

where, ρ is the density, K is a constant, v_p and v_s, respectively, are velocities

of P and S waves, and a periodic function of the forces in time is assumed. Relative amplitude distributions of P and S waves are represented by u_r and (u_θ, u_φ), respectively, on a focal sphere with unit radius. Since u_r is proportional to $2yz$, two P-nodal planes are expressed by $z=0$ and $y=0$, and the direction of the null vector is the x-axis.

Let (ξ, η, ζ) be another right-handed coordinate system related to a geographic coordinate system (**N**, **E**, and **Down**), which may be obtained by suitably rotating the (x, y, z) system. The Cartesian coordinates (ξ, η, ζ) are related to the spherical coordinates (r, Θ, Φ) as follows,

$$\left.\begin{aligned} \xi &= r \sin\Theta \cos\Phi \,, \\ \eta &= r \sin\Theta \sin\Phi \,, \\ \zeta &= r \cos\Theta \,, \end{aligned}\right\} \tag{2}$$

where r is the same as before and Θ refers to the incident angle of the seismic ray at the focus measured from the downward vertical and Φ to the azimuth of the great circle path with respect to the epicenter measured from the north clockwise.

Using direction cosines l_{ik} of the x-, y- and z-axes with respect to the Cartesian coordinates (ξ, η, ζ),

$$\left.\begin{aligned} x &= l_{11}\xi + l_{12}\eta + l_{13}\zeta \,, \\ y &= l_{21}\xi + l_{22}\eta + l_{23}\zeta \,, \\ z &= l_{31}\xi + l_{32}\eta + l_{33}\zeta \,. \end{aligned}\right\} \tag{3}$$

Thus, the problem of determining a fault plane solution is identical with the problem of finding these direction cosines, where independent unknowns are only three because of the orthogonality conditions. By use of these direction cosines, SH and SV components of the S wave motion are written as

$$\left.\begin{aligned} u_\phi &= l_{21}l_{31}a_1 + (l_{22}l_{31} + l_{21}l_{32})a_2 + (l_{23}l_{31} + l_{21}l_{33})a_3 \\ &\quad + (l_{23}l_{32} + l_{22}l_{33})a_4 + l_{23}l_{33}a_5 \,, \\ u_\theta &= l_{21}l_{31}b_1 + (l_{22}l_{31} + l_{21}l_{32})b_2 + (l_{23}l_{31} + l_{21}l_{33})b_3 \\ &\quad + (l_{23}l_{32} + l_{22}l_{33})b_4 + l_{23}l_{33}b_5 \,, \end{aligned}\right\} \tag{4}$$

where, a_j and b_j are known quantities and are expressed in terms of (Θ, Φ) by

$$\left.\begin{aligned} a_1&=-2\sin\Theta\sin 2\Phi\,, & b_1&=\sin 2\Theta\cos 2\Phi\,,\\ a_2&=\sin\Theta\cos 2\Phi\,, & b_2&=(1/2)\sin 2\Theta\sin 2\Phi\,,\\ a_3&=-\cos\Theta\sin\Phi\,, & b_3&=\cos 2\Theta\cos\Phi\,,\\ a_4&=\cos\Theta\cos\Phi\,, & b_4&=\cos 2\Theta\sin\Phi\,,\\ a_5&=-\sin\Theta\sin 2\Phi\,, & b_5&=-\sin 2\Theta\cdot(1+\sin^2\Phi)\,. \end{aligned}\right\} \qquad (5)$$

The angle of polarization of the S wave, γ, is defined theoretically by $\gamma=\tan^{-1}(u_\phi/u_\theta)$. While the observed polarization angle Γ is determined by the angle between the observed S movement and the vertical plane containing the ray. That is, $\Gamma=\tan^{-1}(SH/SV)$, where SH and SV are SH and SV components, respectively, of the S wave amplitude. The sign conventions given by Stauder (1962) are followed in the present paper.

Since the observed angle of polarization is usually defined independently of the sense of the S motion (*cf.* Stauder; 1962), it is convenient to define the residual as the angle between the direction of the theorectical S motion and that of the observed S motion, disregarding the sense of motion. Let R_i be the residual for the i-th station,

$$R_i=\tan^{-1}\left\{\frac{f_i(x_j)-P_ig_i(x_j)}{P_if_i(x_j)+g_i(x_j)}\right\}, \quad j=1,\,2,\,3,\,4 \qquad (6)$$

where

$$\left.\begin{aligned} &P_i=\tan\Gamma_i\,,\\ &f_i(x_j)=u_{\phi i}/(l_{23}l_{33})\,, \quad g_i(x_j)=u_{\theta i}/(l_{23}l_{33})\,,\\ &x_1=l_{31}/l_{33}\,, \quad x_2=l_{32}/l_{33}\,, \quad x_3=l_{21}/l_{23}\,, \quad x_4=l_{22}/l_{23}\,, \end{aligned}\right\} \qquad (7)$$

and R_i is defined in the region $-\pi/2<R_i\leq\pi/2$. The subscript i should be attached also to a_j, b_j, Θ, and Φ in expressions (4) and (5). Since there are only three independent variables, the following condition should be satisfied;

$$x_1x_3+x_2x_4+1=0\,. \qquad (8)$$

If only three observations of the polarization angle are available, fault plane solutions may be obtained directly from the equations,

$$f_i(x_j)-P_ig_i(x_j)=0\,, \quad i=1,\,2,\,3 \qquad (9)$$

under the constraint of equation (8), where the solution is not always determined uniquely as will be seen later.

From the theoretical point of view, it is necessary only to find a minimum of the sum of squares of R_i, that is,

$$\Psi = \sum_{i=1}^{N} R_i^2. \tag{10}$$

3. Numerical Procedure

Because of non-linearity of equations (6), it is not straightforward to find a minimum of Ψ, and usually it is required to have starting values of the unknowns for the iterative process of the numerical method in order to obtain the least squares solution. Moreover, there possibly exist more than one minimum. Therefore, difficulties of the numerical method of the least squares may be concentrated on the following two points.

1. How to find better starting points of the unknown parameters for the successive approximation of the method.
2. How to improve the standard linearization method of the least squares in order to expect better convergency in the iteration and yet not to overlook any minimum around the starting point.

In general, a solution may be obtained graphically by the distribution of the compression-rarefaction of the initial motion of P waves, and the solution can be used as one of the starting points. Sometimes no solution is obtained from the P wave data because of poor spatial distribution of the observing points on the focal sphere. In either case, however, several more starting points are required since it is not assured that there is only one minimum.

If equations (8) and (9) are solved by use of a set of three observations of the polarization angles, three solutions, in general, are obtained. Using many sets of the three observations which are different combinations of the observations from one another, many trial solutions will be obtained. To do this, variables in equations (7) are exchanged for convenience sake.

$$\left.\begin{aligned} K_1 &= x_2 x_4 = -1 - x_1 x_3 , \\ K_2 &= x_1 x_4 + x_2 x_3 , \\ K_3 &= x_1 + x_3 , \\ K_4 &= x_2 + x_4 . \end{aligned}\right\} \tag{11}$$

Thus, f and g in equations (7) are linear functions of K_j, and the condition to be satisfied by K_j is expressed by

$$(K_3K_4-2K_2)^2=(K_4^2-4K_1)\{K_3^2+4(1+K_1)\}\,. \tag{12}$$

If three of the four variables, say, K_1, K_2, and K_3, are eliminated from the three independent linear equations (9) and substituted into equation (12), a cubic equation with respect to K_4 is obtained and, therefore, three roots of the equation may be found in general. The variables x_j are expressed in terms of K_j as follows;

$$\left.\begin{aligned}
&\text{if } K_3K_4-2K_2\geq 0, \text{ then}\\
&\quad x_1=(1/2)\{K_3+\sqrt{K_3^2+4(1+K_1)}\}\,,\\
&\quad x_2=(1/2)\{K_4+\sqrt{K_4^2-4K_1}\}\,,\\
&\quad x_3=(1/2)\{K_3-\sqrt{K_3^2+4(1+K_1)}\}\,,\\
&\quad x_4=(1/2)\{K_4-\sqrt{K_4^2-4K_1}\}\,,\\
&\text{if } K_3K_4-2K_2<0, \text{ then}\\
&\quad x_1=(1/2)\{K_3+\sqrt{K_3^2+4(1+K_1)}\}\,,\\
&\quad x_2=(1/2)\{K_4-\sqrt{K_4^2-4K_1}\}\,,\\
&\quad x_3=(1/2)\{K_3-\sqrt{K_3^2+4(1+K_1)}\}\,,\\
&\quad x_4=(1/2)\{K_4+\sqrt{K_4^2-4K_1}\}\,.
\end{aligned}\right\} \tag{13}$$

Although any combination of three observations taken arbitrarily from the available observations can be used, it may be better to avoid use of the combination of observations from stations located close by one another.

In practice ten to twenty trial solutions are obtained from several sets of the three observations. The sum of squares of the residuals, Ψ, is computed for each of these trial solutions. Several solutions are selected from these trial solutions for the starting points of the parameters in the iterative process of the least squares, for which the values of Ψ are small compared with those for the others. Thus, the first difficulty stated at the beginning of this section may be overcome.

According to the standard method of the least squares, the residual R_i in equation (6) is linearized by omitting the higher order terms in Taylor's expansion, and then the normal equations are obtained. Provided that x_1, x_2, and x_3 in expressions (7) are taken as independent variables, let the initial value of x_j be x_{0j} and the correction of x_j be ε_j in each successive step.

$$\left.\begin{aligned} &x_j = x_{0j} + \varepsilon_j\,, \quad j=1,\,2,\,3 \\ &R_i(x_j) \approx R_i(x_{0j}) + \sum_{j=1}^{3} \varepsilon_j \left[\frac{\partial R_i}{\partial x_j}\right]_{x_j=x_{0j}}. \end{aligned}\right\} \tag{14}$$

The sum of squares of the residuals is approximately given in terms of ε_j by

$$\Psi \approx \Psi_L = \sum_i \left\{ R_i(x_{0j}) + \sum_{j=1}^{3} \left(\varepsilon_j \frac{\partial R_i}{\partial x_{0j}} \right) \right\}^2. \tag{15}$$

Since a minimum of Ψ_L is given by $\dfrac{\partial \Psi_L}{\partial \varepsilon_j} = 0$, the following simultaneous linear equations with respect to ε_j are obtained,

$$\left.\begin{aligned} &a_{11}\varepsilon_1 + a_{12}\varepsilon_2 + a_{13}\varepsilon_3 + a_{01} = 0\,, \\ &a_{12}\varepsilon_1 + a_{22}\varepsilon_2 + a_{23}\varepsilon_3 + a_{02} = 0\,, \\ &a_{13}\varepsilon_1 + a_{23}\varepsilon_2 + a_{33}\varepsilon_3 + a_{03} = 0\,, \end{aligned}\right\} \tag{16}$$

where

$$\left.\begin{aligned} &a_{jk} = \sum_i \left[\frac{\partial R_i}{\partial x_{0j}} \cdot \frac{\partial R_i}{\partial x_{0k}} \right]; \qquad j,\, k \neq 0\,, \\ &a_{0k} = \sum_i \left[R_i \cdot \frac{\partial R_i}{\partial x_{0k}} \right]; \qquad k \neq 0\,. \end{aligned}\right\} \tag{17}$$

Solving the normal equations (16), corrections ε_j for x_{0j} are directly found. If the higher terms of Taylor's expansion are not negligible in expression (15), the minimum point of Ψ_L cannot be taken as an approximate minimum point of Ψ. In such cases it is undesirable to take the values of x_j equal to $x_{0j}+\varepsilon_j$ as a starting point in the next iterative step.

The method of the steepest descent (A. D. Booth; 1955) in solving non-linear simultaneous equations may be applied to this problem. The concept of the method is as follows: $\Psi = const.$ represents a curved surface or an error surface, where Ψ cannot be negative. At a given point (x_{01}, x_{02}, x_{03}) on the error surface, a vector of corrections for x_{0j} should be taken in the direction of the normal to the surface. Let components of the vector of correction be δ_j, which are given by

$$\delta_j = \frac{\partial \Psi}{\partial x_{0j}}\,; \qquad j=1,\,2,\,3\,.$$

Using δ_j, actual corrections of x_j are given by $\alpha\delta_j$, α being a constant.

Although α may be determined in many ways, the simplest method is adopted in the present study in order to make use of the results in equations (17). If appropriate values of α_1 and α_2 are obtained so that $\Psi(x_j)$ may be approximated by a quadratic form with respect to α near the point of x_{0j}, a minimum of the quadratic function at $\alpha=\alpha_0$ is obtained by $\Psi(x_{0j})$, $\Psi(x_{0j}+\alpha_1\delta_j)$, and $\Psi(x_{0j}+\alpha_2\delta_j)$, where the α is regarded temporarily as a variable.

Although α_1 and α_2 may be determined rather arbitrarily if only $|\alpha|$ is sufficiently small, in practice α_1 and α_2 may be given in the following way (*cf.* A. D. Booth; 1955). Using the notations given in (17),

$$\alpha_2=-\Psi(x_{0j})/\left(4\sum_{k=1}^{3}a_{0k}^2\right),$$

$$\alpha_1=\alpha_2/2\,.$$

If the quadratic function given by the above values of α does not have a minimum but a maximum, $(\alpha_2/4,\ \alpha_2/2)$ are used instead of the above $(\alpha_1,\ \alpha_2)$. If the quadratic function constructed by $(\alpha_2/4,\ \alpha_2/2)$ does not have a minimum either, in other words, if $\Psi(x_{0j}+\alpha_2\delta_j/4)$ is larger than $\Psi(x_{0j})$, α_0 is directly taken to be $\alpha_2/80$ by experience, since the value of α_0 is arbitrary and the direction of correction δ_j is correct as far as $|\alpha_0\delta_j|$ is sufficiently small.

Thus two kinds of correction for x_{0j} are obtained, that is, ε_j from the method of linearization and $\alpha_0\delta_j$ from the method of steepest descent. Comparing the value of $\Psi(x_{0j}+\varepsilon_j)$ with that of $\Psi(x_{0j}+\alpha_0\delta_j)$, $(x_{0j}+\varepsilon_j)$ or $(x_{0j}+\alpha_0\delta_j)$ is adopted as a starting point of x_j in the next step.

The method of steepest descent has the advantage of the other in some points. Because by using this method one may confine magnitude of the correction for x_j within small quantity so as to satisfy the condition that $\Psi(x_{0j}+\alpha_0\delta_j)<\Psi(x_{0j})$ and not to overlook any minimum point of Ψ around a given starting point of x_j. This advantage may be very helpful for the present problem where some secondary minima may exist. However, when x_{0j} is located at a point very near a minimum of $\Psi(x_j)$ in a certain successive step, the method is no longer useful especially in the case of less accurate observations. In the present computer program, therefore, the method of linearization only does work when $\sum_j\left(\frac{\partial\Psi}{\partial x_j}\right)^2<10^{-3}$. Thus, the second problem stated in the beginning of this section may be solved.

4. Standard Deviations for Arbitrary Quantities

At present the computer program for the problem consists of the following three parts.

(i) Computing starting points of variables x_j.

(ii) Finding minima, more than one in general, of $\Psi(x_j)$ from the above starting points by the successive method of the least squares.

(iii) Calculating standard deviations for the quantities, dip directions, dip angles, *etc.*, according to the minima obtained above.

Practical processes for (i) and (ii) have already been stated in the previous section. For a given starting point of x_j, the computer program for the successive approximation for (ii) is prescribed to stop the computation at the tenth step of iterations or at the step where $\sum_j \left(\frac{\partial \Psi}{\partial x_j}\right)^2 < 10^{-8}$. In some unfavorable cases it happens that the sum of squares of the partial derivatives becomes almost zero but x_j does not converge to a fixed point. The reason may be apparent if one recalls definitions of Ψ and x_j. The variable x_j can take any value from negative to positive infinity, while the sum of squares of the residuals, Ψ, should be finite for any value of x_j. That is to say, the method may hardly be applied to the case of, for instance, purely vertical strike slip unless the method is modified. To avoid such unfavorable cases, and to confirm the minimum obtained in part (ii) more definitely, the new combination of independent variables x_j' is taken in the program for (iii) instead of x_j in the former part. Using these new variables several steps of the iteration of the linearization method are operated in the process of (iii).

$$\left.\begin{aligned} &x_1'=l_{31}/l_{33}\,, \quad x_2'=l_{32}/l_{33}\,, \quad x_3'=l_{23}/l_{22}\,, \\ &x_4'=l_{21}/l_{22}=-(x_2'+x_3')/x_1'\,, \end{aligned}\right\} \tag{18}$$

$$\left.\begin{aligned} &f_i'(x_j')=u_{\varphi i}/(l_{22}l_{33})\,, \\ &g_i'(x_j')=u_{\theta i}/(l_{22}l_{33})\,. \end{aligned}\right\} \tag{19}$$

Substituting expressions (19) into (6) instead of $f_i(x_j)$ and $g_i(x_j)$, the residual R_i is obtained for the new variables x_j'. At the final stage of the program high accuracy in the numerical calculation may be attained by requiring satisfaction of the condition that $\sum_{j=1}^{3} \varepsilon_j^2 < 10^{-9}$ and $\sum_{j=1}^{3} \left(\frac{\partial \Psi}{\partial x_j'}\right)^2 < 10^{-9}$.

The unbiased estimate of population variance of the observations may

be expressed by

$$\hat{\sigma}^2=\Psi/(N-3), \tag{20}$$

where the observation means deviation of the observed polarization angle from the polarization angle expected theoretically. The estimate of the standard deviation of population of the observations may be given by

$$\hat{\sigma}=\sqrt{\Psi/(N-3)}. \tag{21}$$

The standard deviation thus obtained is used for expressing accuracy of the observation.

It is remarked, however, that the average of a sample of the residuals is not necessarily equal to zero because of non-linearity of the problem. The average of the residuals ($\overline{R}$) may provide a kind of measure for reliability of the solution. Since the population average should be equal to zero, the above definitions of the variance (20) and of the standard deviation (21) will be used.

In order to derive standard deviations of the most probable values for arbitrary functions of x_j', each element of the variance matrix is computed. Let the elements be V_{ij}, where $i, j=1, 2, 3$, and $V_{ij}=V_{ji}$. The weight for the variable x_j' is expressed by $1/V_{jj}$, and $V_{ij}(i\neq j)$ gives the covariance between x_i' and x_j'. The weight p_h for an arbitrary function $h(x_j')$ is given by

$$\frac{1}{p_h}=\sum_{i=1}^{3} V_{ii}\cdot\left(\frac{\partial h}{\partial x_i'}\right)^2+2\sum_{i<j} V_{ij}\cdot\left(\frac{\partial h}{\partial x_i'}\right)\left(\frac{\partial h}{\partial x_j'}\right). \tag{22}$$

The standard deviation of the most probable value of $h(x_j')$ may be given by

$$\hat{\sigma}_h=\hat{\sigma}/\sqrt{p_h}. \tag{23}$$

In the present computer program standard deviations of the following quantities are computed; dip directions and dip angles of the two P-nodal planes, slip angles, trends and plunges of the pressure and tension axes and of the null axis. It is noted that the standard deviation for the trend of an axis has different meaning from that for the plunge. That is, the latter always refers to angular variation along a great circle, whereas the former is along a minor circle and loses its practical meaning as the plunge tends to 90 degrees. It may be convenient, however, to present the standard deviation of the trend in some particular cases, for

example, in the study on the distribution of horizontal pressure directions of the earthquake mechanism.

5. Numerical Examples

As an example the numerical method is applied to an earthquake which took place off the coast of Kamchatka, July 22, 1953, $05^h\ 11^m\ 15^s$, 51° *N*, 157° *E*, $h=60$ km (**USCGS**), the magnitude being $6\,^3/_4$. For this earthquake, Stevens (1964) and Udias (1964) have independently presented numerical solutions from *S* wave data. All the data requisite for the present study, such as the azimuth of the great circle path with respect to the epicenter, the angle of incidence of the ray at the focus, and the observed polarization angle, are given by Udias (1964). For the purpose of comparison of the final results no modification of the given data is made.

Table 1 shows eight sets of observations constructed from 19 observations and 18 trial solutions obtained by using these sets and equation (9).

Table 1. Combinations of observing points and obtained trial solutions.

Stations used	Solution No.	Ψ (in rad^2)	Order of Ψ
(Sol. by Udias)	1	0.64	3
PAL, SLC, SVE	2	5.21	—
CLE, TUC, COP	3	1.40	7
	4	8.45	—
	5	1.19	5
SAS, PAS, STR	6	4.26	—
	7	7.34	—
	8	0.81	4
SLM, RIV, KIR	9	1.37	6
	10	6.21	—
	11	2.48	—
FLO, BAG, DBN	12	4.19	—
	13	6.13	—
	14	0.54	1
BOZ, TAS, KEW	15	13.28	—
SLC, SVE, SCO	16	5.15	—
PAL, BAG, STR	17	5.21	—
	18	12.57	—
	19	0.55	2

The sum of squares of the residuals (Ψ) corresponding to each trial solution is also presented. The first trial solution in the table is the numerical solution given by Udias (1964). Seven trial solutions are selected for the starting values in the successive method of the least squares, which show smaller values of Ψ than the other solutions.

From these starting points all the cases converged to the same minimum and no other minimum was found. At the final stage of the successive approximation,

$$\sum_j \left(\frac{\partial \Psi}{\partial x_j'}\right)^2 = 5.5 \times 10^{-10}, \quad \sum \varepsilon_j^2 = 5.5 \times 10^{-11},$$

$$\Psi = 0.46551 \ (\text{rad}^2), \qquad \overline{R} = -0°.6,$$

and unbiased estimate of population variance of the observation (polarization angle) is 0.029094 in square of radian, and estimated standard deviation of the observation is 9°.8. Table 2 shows comparison of the

Table 2. Comparison of the present least squares solution with the solutions given by Stevens (1964) and by Udias and Stauder (1964).

		L. S. Sol.	Stevens (1964)	Udias (1964)
Plane (a)	Dip Direc.	−49°.4 ±4°.4	−44°	−50°.5 ±2°.9
	Dip Angle	41.0 1.1	39	42.7 ±0.4
Plane (b)	Dip Direc.	138.9 6.4	142	143.6
	Dip Angle	49.3 1.2	51	48.9
Pressure Axis	Trend	135.1 5.1	140	136.9
	Plunge	4.2 1.1	6	3.1
Tension Axis	Trend	0.0 19.1	−15	23.4
	Plunge	84.1 2.0	83	82.2
Null Axis	Trend	−134.6 5.4	−130	−132.7
	Plunge	4.1 8.0	3	7.1
Slip Angle	Plane (a)	83.7 3.7	85	75.7
	Plane (b)	84.5 6.9	86	77.2

present solution with the results given by Stevens (1964) and by Udias and Stauder (1964). It may be reasonable that the quantities with small standard deviation of the present solution agree well with one another among the three solutions. Relatively small standard deviation of the polarization angles suggests that the data for the Kamchatka earthquake are well refined, and actually there was little trouble in obtaining the

least squares solution by the present method.

There are many examples of earthquakes in which from some of the selected starting points the successive approximations do not converge. As an example of such cases an earthquake which took place near Japan, October 8, 1960, may be suitable. For this earthquake the data of the polarization angle have been given by Ritsema (1965) and detailed discussions will be made in the second paper. A graphical solution given by Ritsema is taken as one of the starting points and the sum of squares of the residuals (Ψ) is equal to 6.50 at the starting point. Using 15 observations, four starting points are chosen besides the above graphical solution. From the graphical solution and from two of the four trial solutions obtained numerically, for which $\Psi=1.51$ and 3.15, respectively, the successive approximations converge to the same minimum point at which $\Psi_m=0.8539$. From one of the other trial solutions for which $\Psi=2.88$ the approximation fails to converge. From the last trial solution ($\Psi=3.99$) the method converges to another minimum point at which $\Psi_m=3.539$. For the purpose of comparison, major quantities for the two tentative solutions are tabulated in Tables 3 a and 3 b.

Table 3 a.

	Ψ	$\Sigma\left(\frac{\partial\Psi}{\partial x_j'}\right)^2$	$\Sigma\varepsilon_j^2$	$\hat{\sigma}^2$ (rad^2)	$\hat{\sigma}$	$\overline{R}$
Sol. (1)	0.854	7.4×10^{-10}	6.0×10^{-15}	0.0712	15°.3	4°.4
Sol. (2)	3.54	4.7×10^{-10}	1.3×10^{-12}	0.2949	31°.1	3°.8

Table 3 b.

	Dip Direc.	Dip Angle	Dip Direc.	Dip Angle
Sol. (1)	116°.8±28°.1	10°.5±1°.1	−61°.9±5°.2	79°.5±1°.1
Sol. (2)	77°.1± 4°.7	42°.9±9°.0	−20°.9±5°.7	82°.6±4°.0

The difference of the variances of the observations between the two solutions is so big that it may be natural to choose the solution (1) as the most probable solution from the S wave data, while standard deviations for dip directions and dip angles of the P-nodal planes as shown in Table 3 b

are rather in favor of the solution (2) as a whole. This circumstance may be suggestive. It is naturally concluded that one should not discuss for superiority of the solutions by means of standard deviations for such quantities as dip directions and dip angles so far as the normal distribution of population of the observation (polarization angle) is assumed.

Generally speaking, significance of the difference of the variances should be subjected to the statistical test. This may be done by testing the null hypothesis that the population variances of the observations are equal to each other by use of the F-distribution where five per cent of the level of significance is taken in the present study. If the null hypothesis is rejected by the test, one may take the solution which has smaller variance than the other as the most probable solution with the five per cent of the level of significance. If it turns out to be no significant difference of the variances by the test, both of the two solutions should be adopted tentatively as far as the S wave data is concerned, and the best solution should be determined in contrast with the observations of P wave.

For the present earthquake the statistical test shows that the difference of the two variances is significant. Thus, solution (1) is the most probable solution determined indepedently of the P wave data. As will be seen in the second paper, the solution (1) is a better solution for the P wave data than the other. This example may suggest that the use of preliminary solution from the P wave data is desirable as one of starting points.

6. Summary and Acknowledgements

Assuming that a sample of the residuals, deviations of the observed polarization angles from those expected for the theoretical model of the double couple hypothesis, is extracted from a population of the normal distribution, a least squares method is presented in which the sum of squares of the residuals is minimized.

In practice starting points of successive approximations in the method of least squares are obtained by solving directly non-linear equations with respect to three independent variables which are constructed by three observations. Two different kinds of numerical method are co-operated in the computer program of the successive approximation process, and the program is so carefully designed that any minimum around the starting point may not be overlooked.

Standard deviations of such quantities as dip direction, dip angle, *etc.*

which should be obtained in the focal mechanism study are computed. These standard deviations depend upon variance of the observations and size of the sample and largely on the spatial distribution of observing points on the focal sphere. For any quantity expressed by a function of the parameters (x_1', x_2', x_3') of the estimated solution the standard deviation can be obtained by using numerical values of elements of the variance matrix.

The method is successfully applied to two examples of earthquakes. For one of the two examples two minimum points were found. For such cases, in general, use of statistical test is proposed in order to distinguish the most probable solution from the two solutions.

Thus, the fault plane solution of an earthquake is obtainable independently of the P wave data by use of the analytical method in which any preliminary solution is not necessarily required.

The author is greatly indebted to Rev. William Stauder, S. J. who suggested the problem of the focal mechanism determination and offered valuable advice. He wishes to thank Prof. H. Honda and Prof. T. Asada for their discussions and encouragements. He is also grateful to Dr. R. Sato who read the paper in manuscript and provided many helpful comments.

References

ADAMS, W. M.
1958, "A Study of Earthquake Mechanism Using *S*-Wave Data", *Bull. Seism. Soc. Amer.*, **48**, 201–219.

BOOTH, A. D.
1955, "*Numerical Methods*", London; Butterworths Scientific Publ.

HODGSON, J. H. and W. M. ADAMS
1958, "A Study of the Inconsistent Observations in the Fault-Plane Project", *Bull. Seism. Soc. Amer.*, **48**, 17–31.

HODGSON, J. H. and A. J. WICKENS
1965, "Computer-Determined *P*-Nodal Solutions for the Larger Earthquakes of 1959–1962", *Publ. Dom. Obs., Ottawa*, **31**, 123–143.

HONDA, H.
1962, "Earthquake Mechanism and Seismic Waves", *Geophysical Notes, Tokyo Univ.*, **15**, *Supplement*, 1–97.

KASAHARA, K.
1963, "Computer Program for a Fault-Plane Solution", *Bull. Seism. Soc. Amer.*, **53**, 1–13.

KNOPOFF, L.
1961*a*, "Analytical Calculation of the Fault-Plane Problem", *Publ. Dom. Obs.*, **24**, 309–315.

KNOPOFF, L.
1961*b*, "Statistical Accuracy of the Fault-Plane Problem", *Publ. Dom. Obs.*, **24**, 317–319.

RITSEMA, A. R.
1965, "The Mechanism of Some Deep and Intermediate Earthquakes in the Region of Japan", *Bull. Earthq. Res. Inst.*, **43**, 39-52.
STAUDER, W., S. J.
1962, "The Focal Mechanism of Earthquakes", *Advances in Geophysics*, **9**, 1-76.
STAUDER, W., S. J. and G. A. BOLLINGER
1964, "The *S*-Wave Project for Focal Mechanism Studies: Earthquakes of 1962", *Bull. Seism. Soc. Amer.*, **54**, 2199-2208.
STEVENS, A. E.
1964, "Earthquake Mechanism Determination by *S*-Wave Data", *Bull. Seism. Soc. Amer.*, **54**, 457-474.
UDIAS, A., S. J.
1964, "A Least Squares Method for Earthquake Mechanism Determination Using *S*-Wave Data", *Bull. Seism. Soc. Amer.*, **54**, 2037-2047.
UDIAS, A., S. J. and W. STAUDER, S. J.
1964, "Application of Numerical Method for *S*-Wave Focal Mechanism Determinations to Earthquakes of Kamchatka-Kurile Islands Region", *Bull. Seism. Soc. Amer.*, **54**, 2049-2065.

---

47. *S* 波観測から地震のメカニズムを求める最小二乗法：第一報

東京大学理学部・地球物理学教室　平　沢　朋　郎

S 波の polarization angle の観測値と double couple の仮説に基く理論値の差（残差）を確率変数とみなし，残差平方和を最小にするメカニズムの解を求める解析的方法が与えられた．

問題が非線型であるため，残差平方和の極小点は一つとは限らない．従って最小二乗解を求める逐次近似法には数ケの出発点を求める事が望ましい．それらの出発値は三ケの観測値の任意の組合わせを用いて直接メカニズムの解を求める事により得られる．すなわち，この方法に対してはメカニズムの解に関する図式解その他，何んらの予備知識も要求されない．また，二ケ以上の極小点が見出された場合には，分散の差の有意性を検定する事によって *P* 波の観測とは独立な解を得る事ができる．

得られた最小二乗解から計算される，いかなる物理量に対しても，その標準偏差は独立変数の分散および共分散を使って計算できるが，プログラムの中で計算される量は，二つの *P* 波節面の dip direction と dip angle, null axis, pressure axis, および tension axis の trend と plunge, それに slip angle である．

なお，二つの地震に対してその数値例が示されている．