GRAPHICAL EXPRESSION OF THE LINEAR LEAST-SQUARE THEORY, AND ITS APPLICATION TO THE VLBI OBSERVATIONAL EQUATIONS

Koichi Nakajima

I. *Introduction*

In the geodetic observations such as Very Long Baseline Interferometry (VLBI) the method of least squares is generally used for determining geodetic parameters from observed data. In most cases the observational equations are linearized, and its theoretical treatment is given as the theory of linear estimation in which the Gauss-Markoff model is applied (Rao 1973).

The linear least-square theory (we abbreviate LLST hereafter) is described in terms of vectors and matrices. It seems that some parts should be expressed graphically to help visual understandings. However, systematic explanations utilizing graphical expressions are hardly seen.

Here we try to give some simplified graphical explanations of a part of LLST, to make clear the characteristics of the normal equation. Then we apply this analysis to an example, a VLBI observation schedule.

The summary of the method of analysis is as follows. The coefficient matrix of the normal equation determines a (hyper)ellipsoid in an m-dimensional space formed by m unknown parameters, and if the composition of the observational equations is improper, some axes are elongated abnormally, causing larger errors in many parameter estimations. This is because the standard deviation of each parameter is given as the projection of the ellipsoid onto each axis of coordinate corresponding to the parameter. The case like this is often called "bad separation of parameters."

In this case, however, we can find a linear combination of unknown parameters whose s.d. is rather small, choosing the estimation-axis toward a thinner part of the ellipsoid. To obtain the lengths and directions of principal axes of the ellipsoid, we solve an eigen value problem of the coefficient matrix of the normal equation.

Applying this analysis to actual observations, we can judge, quantitatively, the efficiency of the observation schedule, or we can find an optimum estimation of parameters in a form of their linear combination in case that the condition of observation is restricted.

II. *Equations*

In this section equations describing LLST are summarized in terms of vectors and

matrices.   Notations are used following Rao's textbook (Rao 1973).

Let $\mathbf{A}'$ be the transpose of a matrix (including a vector) $\mathbf{A}$.   Then we use following notations:

$\mathbf{Y}'=(y_1, \ldots, y_n)$: observed data such as delay times in VLBI.
  n denotes the number of observations.

$\mathbf{X}=(x_{ij} \mid i=1, \ldots, n; j=1, \ldots, m)$
  : $n \times m$ known matrix of coefficients of observational equations.
  m denotes the number of unknown parameters.

$\boldsymbol{\beta}'=(\beta_1, \ldots, \beta_m)$: unknown parameters.
$\boldsymbol{\varepsilon}'=(\varepsilon_1, \ldots, \varepsilon_n) \equiv (\mathbf{Y}-\mathbf{X}\boldsymbol{\beta})'$: residuals.

Note that each one $\boldsymbol{\beta}$ determines one $\boldsymbol{\varepsilon}$.   We expect that a true $\boldsymbol{\beta}$ determines $\boldsymbol{\varepsilon}$ such as

$E(\boldsymbol{\varepsilon})=0,$

where $E(\boldsymbol{\varepsilon})$ means the expectation of $\boldsymbol{\varepsilon}$.

For simplicity we assume that the observations have equal weight and that y_i 's are independent, i.e.,

$$D(\boldsymbol{\varepsilon})=\sigma^2\mathbf{I}, \ (\sigma>0), \quad \dotfill (1)$$

where $D(\boldsymbol{\varepsilon})=E(\boldsymbol{\varepsilon} \ \boldsymbol{\varepsilon}')$ denotes the variance-covariance of $\boldsymbol{\varepsilon}$, and $\mathbf{I}$ is the unit matrix of $n \times n$. Further we assume each ε_i has the normal distribution

$$N(\varepsilon_i \mid 0, \sigma^2) \equiv (\sigma \sqrt{2\pi})^{-1} \exp(-\varepsilon_i^2/2\sigma^2). \quad \dotfill (2)$$

The observational equations are

$$\mathbf{Y}=\mathbf{X}\boldsymbol{\beta}+\boldsymbol{\varepsilon}, \quad \dotfill (3)$$

which we depict in Figure 1.   In the case of $n \geqq m$, the normal equation is

$$\mathbf{X}'\mathbf{X}\boldsymbol{\beta}=\mathbf{X}'\mathbf{Y}, \quad \dotfill (4)$$
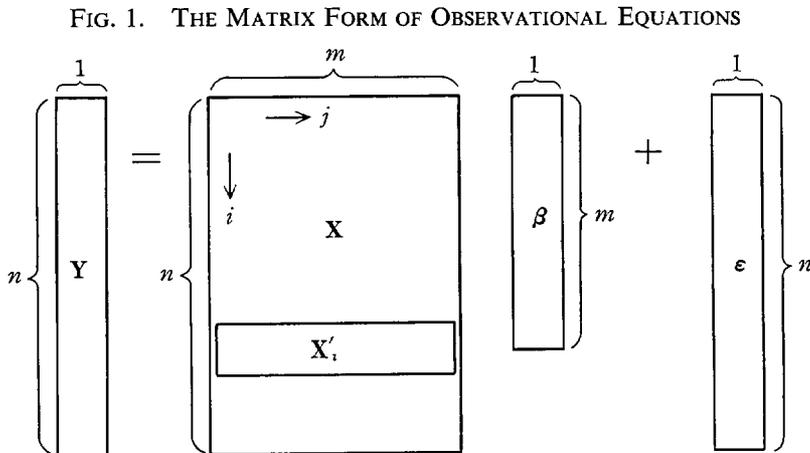
FIG. 1.   THE MATRIX FORM OF OBSERVATIONAL EQUATIONS

which is derived from equation (3) by minimizing $\varepsilon'\varepsilon$, i.e. $(Y-X\beta)'(Y-X\beta)$. Let $\hat{\beta}$ be a solution of (4), and let

$$\hat{\varepsilon} \equiv Y - X\hat{\beta}. \quad \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots (5)$$

Then we have

$$X'\hat{\varepsilon} = X'(Y-X\hat{\beta}) = X'Y - X'X\hat{\beta} = 0. \quad \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots (6)$$

If $\mathrm{rank}X'X = m$ (i.e. $X'X$ is regular) the solution $\hat{\beta}$ is determined uniquely and is given by

$$\hat{\beta} = (X'X)^{-1}X'Y. \quad \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots (7)$$

If it is not the case (i.e. $\mathrm{rank}X'X < m$) we can obtain a $\hat{\beta}$ by using the generalized inverse of matrix (Rao 1973, chapter 4; Bjerhammer 1973). However, we deal with only the former case because in this work we are investigating the efficiency of given observational equations, assuming that $X'X$ is regular.

Note that if $n=m$ and X is regular, we can find a solution $\hat{\beta} = X^{-1}Y$, where $\varepsilon'\varepsilon = 0$. This is a special case of equation (6) because if X is an $m \times m$ matrix and is regular, then $(X'X^{-1} = X^{-1}(X')^{-1}$.

Here we will show that $(X'X)^{-1}$ gives the variance-covariance of estimated $\hat{\beta}$ (i.e. variances and covariances among $\hat{\beta}_1, \dots, \hat{\beta}_m$). From equations (3) and (1) we have

$$D(Y) = D(\varepsilon) = \sigma^2 I.$$
$$\therefore D(X'Y) = E(X'YY'X) = X'E(YY')X = \sigma^2 X'X.$$
$$\therefore D(\hat{\beta}) = D((X'X)^{-1}X'Y) = (X'X)^{-1}D(X'Y)[(X'X)^{-1}]'$$
$$= \sigma^2(X'X)^{-1}(X'X)[(X'X)^{-1}]'$$
$$= \sigma^2(X'X)^{-1}, \quad \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots (8)$$

because $X'X$ and $(X'X)^{-1}$ are symmetric.

III. *Graphical Expressions of LLST and the Error Ellipsoid*

In this section we display graphically the relations among the observational equations, the estimated $\hat{\beta}$, and its error ellipsoid.

III.1. The relation between observational equations and $\hat{\beta}$
In the m-dimensional space of β, an equation of β

$$X_i'\beta = y_i \quad \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots (9)$$

determines a hyperplane, where

$$X_i' \equiv (x_{i_1}, \dots, x_{im}), \quad i = 1, \dots, n.$$

is a row vector of X (see Figure 1). Let

$$\|X_i\| \equiv (X_i'X_i)^{1/2}, \quad \text{and}$$
$$e_i \equiv X_i/\|X_i\| \quad \text{(i.e. a unit vector toward } X_i\text{)}.$$

The hyperplane (9) is perpendicular to e_i (i.e. X_i), and its distance from the origin (i.e. $\beta = 0$)

FIG. 2. RELATIONS AMONG THE HYPERPLANE $X_i'\beta = y_i$, e_i, y_i, δ_i, AND ε_i
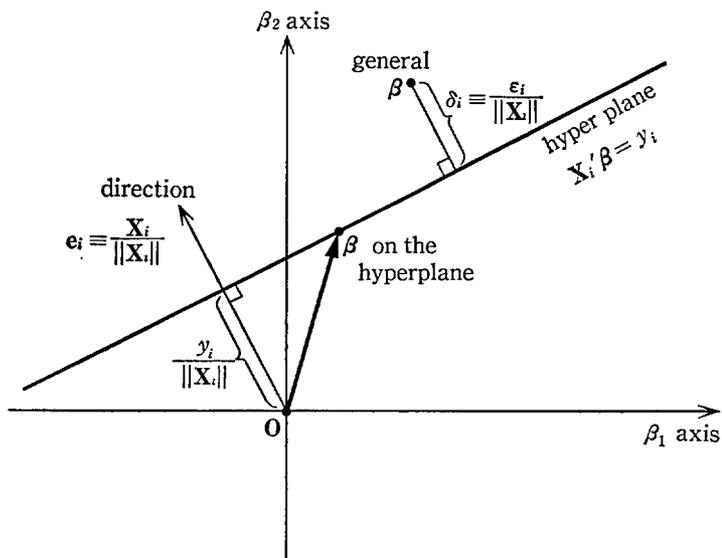
FIG. 3. RELATIONS AMONG THE HYPERPLANES AND $\hat{\beta}$ ESTIMATED BY THE METHOD OF LEAST SQUARES
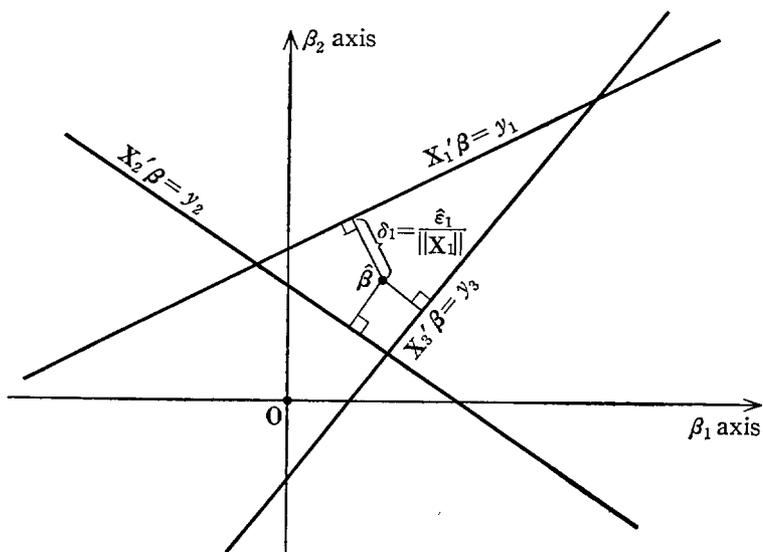
is given by $y_i/\|X_i\|$. This is understood from that the projection of every vector β of (9) onto the direction e_i is given by $e_i'\beta = X_i'\beta/\|X_i\|$ (see Figure 2).

If a point β is not on the plane (9), there is an ε_i which is given by equation (3). In this case the distance between β and the plane is

$$\delta_i \equiv \text{abs}(\varepsilon_i)/\|\mathbf{X}_i\|,$$

as shown in Figure 2.

As seen in Figure 3, the least-square solution $\hat{\beta}$ is in the midst of hyperplanes $\mathbf{X}\beta = \mathbf{Y}$, and is given as a point which makes the square-sum of $\varepsilon_i = \delta_i \|\mathbf{X}_i\|$ (and not that of distances δ_i) be minimum.

III.2.   Distributions of the true β around the estimated $\hat{\beta}$, and the error ellipsoid

Next we investigate the distribution of the true β around the estimated $\hat{\beta}$.   The distribution can be regarded the distribution of $\hat{\beta}$ itself.   From this distribution we will derive the error ellipsoid.

Adopting the Gauss-Markoff model, we have that the distribution of the true β determined by an observational equation

$$\mathbf{X}_i'\beta = y_i + \varepsilon_i \text{ (in which } E(\varepsilon_i^2) = \sigma^2),$$

is given by a one-dimensional normal distribution along $\mathbf{e}_i$, such as

$$
\begin{aligned}
N(\varepsilon_i/\|\mathbf{X}_i\| \mid 0, \, (\sigma/\|\mathbf{X}_i\|)^2) \\
= N((\mathbf{X}_i'\beta - y_i)/\|\mathbf{X}_i\| \mid 0, \, (\sigma/\|\mathbf{X}_i\|)^2) \\
= N(\mathbf{e}_i'(\beta - \bar{\beta}_i) \mid 0, \, (\sigma/\|\mathbf{X}_i\|)^2) \\
= N(\mathbf{e}_i'\beta \mid \mathbf{e}_i'\bar{\beta}_i, \, (\sigma/\|\mathbf{X}_i\|)^2), \quad\quad\quad\quad\quad\quad\quad\quad
\end{aligned}
\tag{10}
$$

where we put $y_i = \mathbf{X}_i'\bar{\beta}_i$ (see Figure 4).

Now let us define the m-dimensional normal distribution by

$$N_m(\beta|\bar{\beta}, \mathbf{v}) = N_m(\bar{\beta}|\bar{\beta}, \mathbf{v}) \exp[-(1/2)(\beta - \bar{\beta})'\mathbf{v}^{-1}(\beta - \bar{\beta})], \quad\quad\quad\quad\quad\quad \tag{11}$$

FIG. 4.   ONE DIMENSIONAL NORMAL DISTRIBUTION ALONG $\mathbf{e}_i$
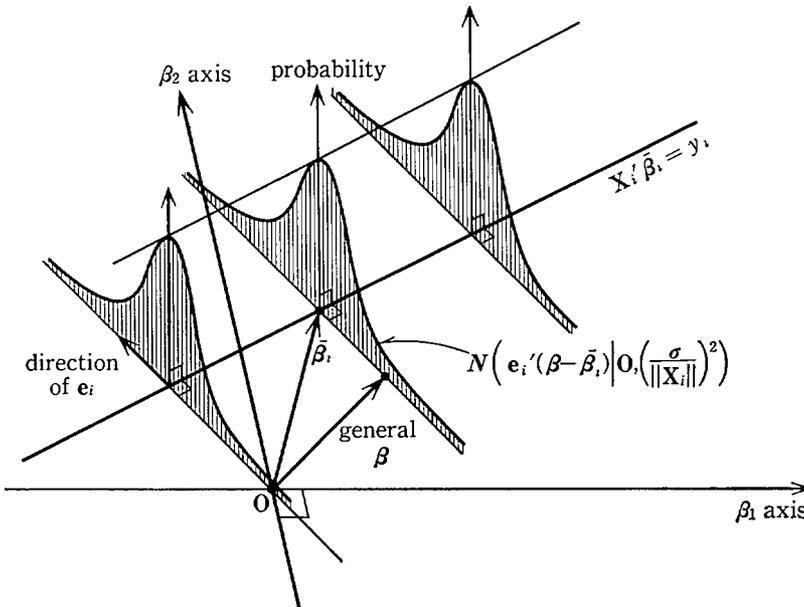
where $\mathbf{v}$ denotes the variance-covariance matrix, and $N_m(\bar{\beta}|\bar{\beta}, \mathbf{v})$ is determined by the normalization, so that $\int N_m(\beta|\bar{\beta}, \mathbf{v})d\beta = 1$. Using these notations, we test whether or not the distribution of the true β around the estimated $\hat{\beta}$ is an m-dimensional normal distribution.

This distribution is calculated by multiplying n one-dimensional normal distributions of equation (10), in the m-dimensional space of β. Rewriting equation (10) into m-dimensional form, as

$$N_m(\beta|\bar{\beta}_i, \mathbf{v}_i)$$
$$= N_m(\bar{\beta}_i|\bar{\beta}_i, \mathbf{v}_i) \exp[-(1/2)\{e_i'(\beta-\bar{\beta}_i)\}^2/(\sigma/||\mathbf{X}_i||)^2]$$
$$= N_m(\bar{\beta}_i|\bar{\beta}_i, \mathbf{v}_i) \exp[-(1/2)\{\mathbf{X}_i'(\beta-\bar{\beta}_i)\}^2/\sigma^2], \quad \dots\dots\dots\dots\dots\dots(12)$$

we have

$$\prod_{i=1}^{n} N_m(\beta|\bar{\beta}_i, \mathbf{v}_i)$$
$$= [\prod N_m(\bar{\beta}_i|\bar{\beta}_i, \mathbf{v}_i)] \exp[-(1/2\sigma^2) \sum_{i=1}^{n} \{\mathbf{X}_i'(\beta-\bar{\beta}_i)\}^2]. \quad \dots\dots\dots\dots\dots(13)$$

Remembering that $\mathbf{X}_i'\bar{\beta}_i = y_i = \mathbf{X}_i'\hat{\beta} - \hat{\varepsilon}_i$ from equation (5), we have

$$\sum_{i=1}^{n} \{\mathbf{X}_i'(\beta-\bar{\beta}_i)\}^2 = \sum \{\mathbf{X}_i'(\beta-\hat{\beta}_i) + \varepsilon_i\}^2$$
$$= \sum \{(\beta-\hat{\beta}_i)'\mathbf{X}_i\mathbf{X}_i'(\beta-\hat{\beta}_i) + 2\hat{\varepsilon}_i\mathbf{X}_i'(\beta-\hat{\beta}_i) + \hat{\varepsilon}_i^2\}$$
$$= (\beta-\hat{\beta})'\mathbf{X}'\mathbf{X}(\beta-\hat{\beta}) + 2\hat{\varepsilon}'\mathbf{X}(\beta-\hat{\beta}) + \hat{\varepsilon}'\hat{\varepsilon}. \quad \dots\dots\dots\dots\dots\dots(14)$$

From equation (6) the second term drops, and we have

$$\prod_{i=1}^{n} N_m(\beta|\bar{\beta}_i, \mathbf{v}_i)$$
$$= [\prod N_m(\bar{\beta}_i|\bar{\beta}_i, \mathbf{v}_i)] \exp[-(1/2\sigma^2)(\beta-\hat{\beta})'\mathbf{X}'\mathbf{X}(\beta-\hat{\beta})] \exp[-(1/2\sigma^2)\hat{\varepsilon}'\hat{\varepsilon}]. \quad \dots\dots(15)$$

FIG. 5.   RELATIONS AMONG EACH HYPERPLANE AND THE SDE
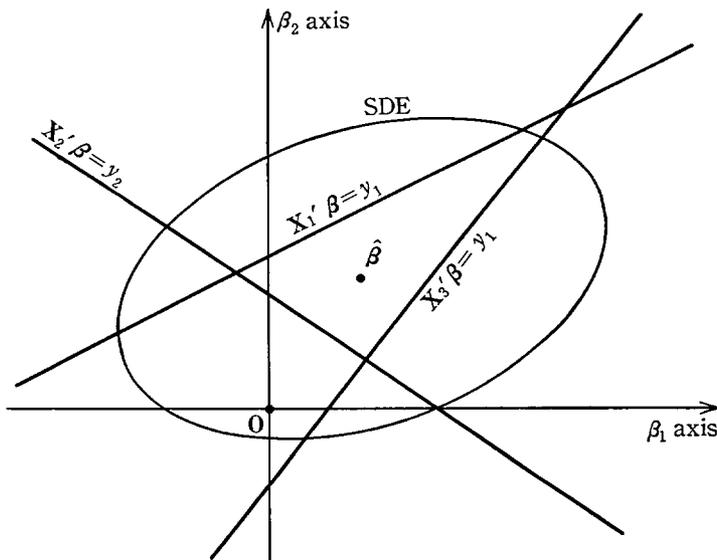
Thus we have proved that this distribution is an m-dimensional normal distribution whose variance-covariance matrix is $\sigma^2(X'X)^{-1}$. Also we have shown equation (8) again. It can be seen in equation (15) that the peak of distribution is the highest when $\hat{\beta}$ is so determined that $\hat{\varepsilon}'\hat{\varepsilon}$ is minimum.

From these analyses we find that those β which give the probability of $\exp(-1/2)$ of the peak value form a hyperellipsoid in the m-dimensional space, such as

$$\sigma^2 = (\beta - \hat{\beta})'X'X(\beta - \hat{\beta}). \quad\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots(16)$$

The right hand side of equation (16) shows a quadratic form of β which is positive except $\beta = \hat{\beta}$, and then the symmetric matrix $X'X$ is positive definit (see Rao 1973, 1c). Thus this ellipsoid is everywhere convex. Let us call it a "standard-deviation ellipsoid" (SDE) because such a $\beta - \hat{\beta}$ is called a standard-deviation. In Figure 5 we show graphically the relations among each hyperplane for each observational equation and the SDE.

III.3.   The properties of the SDE

In this subsection we investigate the properties of the SDE and their relations to the parameters which appear in the LLST. The properties of an ellipsoid are mainly the lengths and directions of its principal axes, and to find them is an eigen-value problem. For simplicity we put $\hat{\beta} = O$ hereafter.

In this case equation (16) becomes

$$\sigma^2 = \beta'X'X\beta. \quad\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots(17)$$

Since $X'X$ is a positive-definit $m \times m$ matrix as described above, $X'X$ has m eigen values $\lambda_j(j = 1, \dots, m)$ which are all positive. In this case we have an orthogonal $m \times m$ matrix

FIG. 6.   RELATIONS AMONG β_j, P_j, γ_j, λ_j, AND THE SDE
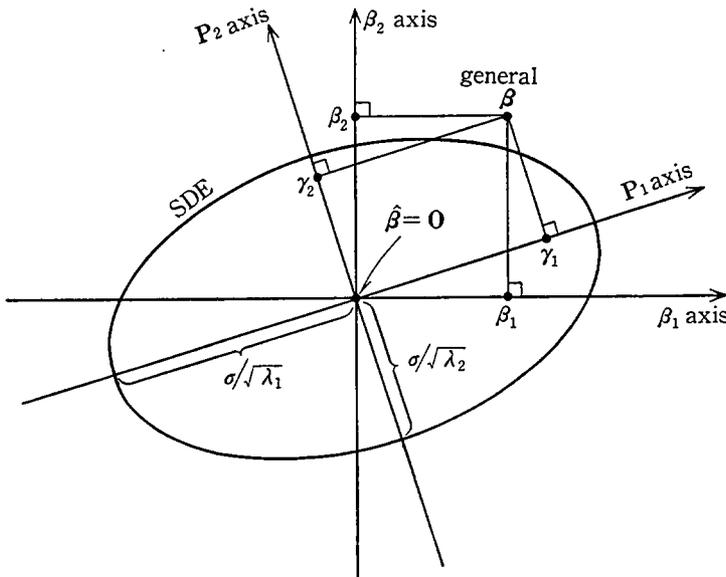
P, such as

$$|\mathbf{P}| \equiv (\text{determinant of } \mathbf{P}) = 1, \text{ and}$$

$$\gamma \equiv \mathbf{P}'\beta \quad \text{or} \quad \beta = \mathbf{P}\gamma, \quad \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots(18)$$

and

$$\beta'\mathbf{X}'\mathbf{X}\beta = \gamma'\mathbf{P}'\mathbf{X}'\mathbf{X}\,\mathbf{P}\gamma$$
$$= \gamma'\begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_m \end{pmatrix}\gamma$$
$$= \sum_{j=1}^{m} \lambda_j \gamma_j^2. \quad \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots(19)$$

The columnal vectors $\mathbf{P}_j$ $(j=1,\dots,m)$ of $\mathbf{P}$ are the eigen vectors which are shown in Figure 6. In this case, $\gamma'=(\gamma_1,\dots,\gamma_m)$ expresses a set of coordinates of a point β with respect to the reference framework formed by $\mathbf{P}_1,\dots,\mathbf{P}_m$.

From equations (17) and (19) we have

$$\sigma^2 = \sum_{j=1}^{m} \lambda_j \gamma_j^2, \quad \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots(20)$$

which determines a hyperellipsoid whose principal axis along $\mathbf{P}_j$ has length of $\sigma\sqrt{\lambda_j}$ (see Figure 6).

Let us examine the relation between the SDE and each s.d. of the estimated value of an unknown parameter β_j. Calculating the distribution of the true β_j, we will find its s.d. as a function of parameters which determine the SDE. Note that from equation (8) we already know that the variance of a true β_j (or $\hat\beta_j$) is given by the jj component of $\sigma^2(\mathbf{X}'\mathbf{X})^{-1}$.

The distribution of β_j is given by

$$\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} N_m(\beta|0,\,v)d\beta_1,\dots\check{\jmath}\dots,d\beta_m. \quad \dots\dots\dots\dots\dots\dots(21)$$

To calculate one integral $\int_{-\infty}^{\infty} N_m(\beta|0,\,v)d\beta_k(k\neq j)$, let us see the β_k-dependence of $N_m(\beta|0,\,v)$ which we denote $f_m(\beta_k)$. Since the exponent in $N_m(\beta|0,\,v)$ is a quadratic form of β_k, we can write

$$f_m(\beta_k) \propto \exp[-(1/2\sigma^2)\{a^2(\beta_k-b)^2+c\}], \quad \dots\dots\dots\dots\dots\dots\dots(22)$$

where $a^2=\sum_{i=1}^{m} x_{ik}^2$, and b is a linear combination of $\beta_1,\dots\check{k}\dots,\beta_m$, further c is a quadratic form of $\beta_1,\dots\check{k}\dots,\beta_m$.

Since a is independent of β, the integral $\int N_m(\beta|0,\,v)d\beta_k=\int f_m(\beta_k)d\beta_k$ is proportional to the peak value of $f_m(\beta_k)$. (The coefficient of the proportion is $\sigma\sqrt{\pi}/a$.) Because of the same reason, the integral (21) is proportional to the peak value of $N_m(\beta|0,\,v)$ as a function of $\beta_1,\dots\check{k}\dots,\beta_m$, that is, (21) is proportional to the peak value of $N_m(\beta|0,\,v)$ on the hyperplane given by $\beta_j=$constant (see Figure 7). Since the coefficient of proportion is independent of β_j, the distribution of $\hat\beta_j$ is "normal" about $\beta_j=0$.

Then it is easy to find the s.d. of β_j, i.e. a β_j at which the probability is $\exp(-1/2)$ of the peak value that occurs at $\beta_j=0$, as follows. Such a β_j is given as the coordinates β_{j+}

and β_{j-} of the two extreme points on the SDE toward the direction of β_j axis (see Figure 7), because at these points the distribution $N_m(\beta|0, v)$ confined in the hyperplane $\beta_j = \beta_{j+}$ or $\beta_j = \beta_{j-}$ is maximum, and is $\exp(-1/2)$ of the peak value of $N_m(\beta|0, v)$ which occurs at $\beta = 0$.

Finally let us see that the jj component of $\sigma^2(X'X)^{-1}$ (i.e. the variance of $\hat{\beta}_j$) surely corresponds to $(\beta_{j+})^2$ or $(\beta_{j-})^2$. To begin with, let us calculate the m-dimensional "volume" of the ellipsoid $\sigma^2 = \beta'X'X\beta$. From equation (20) and from Figure 6 we see that the volume V_p of the parallelotope which surrounds the SDE is given by

$$V_p = \prod_{j=1}^{m} \sigma/\sqrt{\lambda_j}. \qquad \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots(23)$$

In other words the determinant of $\sigma^2(X'X)^{-1}$ is, from equation (19),

$$\begin{aligned}
\sigma^{2m}|(X'X)^{-1}| &= \sigma^{2m}|X'X|^{-1}\\
&= \sigma^{2m}|P'X'XP|^{-1}\\
&= \sigma^{2m}(\prod_{j=1}^{m} \lambda_j)^{-1}\\
&= V_p^2, \qquad \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots(24)
\end{aligned}$$

because we defined $|P| = 1$. The ratio of the volume of the SDE (denoted by V_e) to that of the parallelotope (V_p) is equal to the ratio of the volume of an m-dimensional hypersphere to that of a hypercube, and is equal to

$$K_m = \pi^{m/2}/\Gamma(m/2+1), \qquad \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots(24)$$

that is,

$$V_e = K_m \sigma^m \sqrt{abs |X'X|^{-1}}. \qquad \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots(25)$$

Next we calculate the "area" of a cross-section of the SDE, cut by a hyperplane $\beta_j = 0$. The

FIG. 7. RELATION BETWEEN THE STANDARD DEVIATION OF $\hat{\beta}_j$ AND THE SDE
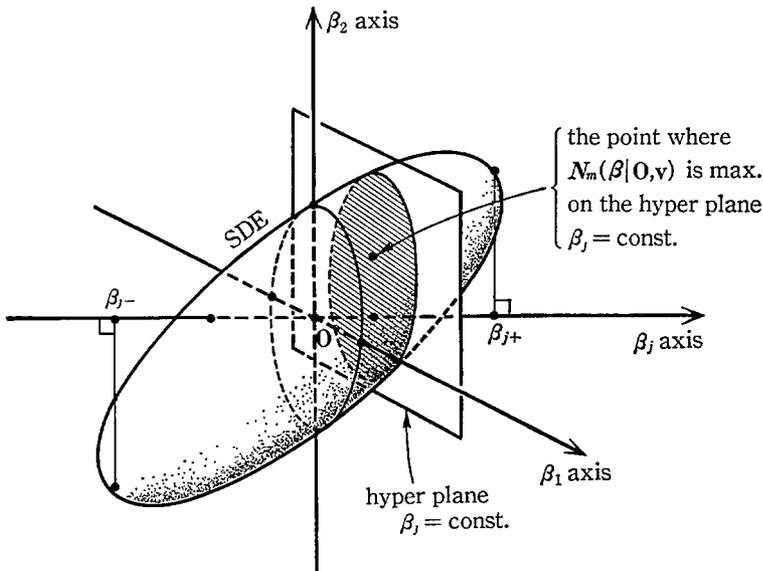
cross-section is an $(m-1)$-dimensional hyperellipsoid which is given from equation (17) by putting $\beta_j=0$, and is

$$\sigma^2=(\beta_1, \ldots, 0, \ldots, \beta_m)\mathbf{X'X}\begin{pmatrix} \beta_1 \\ \vdots \\ 0 \\ \vdots \\ \beta_m \end{pmatrix},$$

that is equal to

$$\sigma^2=(\beta_1, \ldots, \overset{\vee}{j} \ldots, \beta_m)\mathbf{S}\begin{pmatrix} \beta_1 \\ \vdots \\ j< \\ \vdots \\ \beta_m \end{pmatrix}, \dots\dots\dots\dots\dots\dots\dots\dots\dots(26)$$

where S denotes an $(m-1)\times(m-1)$ matrix which is made from $\mathbf{X'X}$ by omitting j-th row and j-th column.   Therefore, the area of the cross-section (denoted by S_e) is given by

$$S_e=K_{m-1}\sigma^{m-1}\ \sqrt{\mathrm{abs}\ |\mathbf{S}|^{-1}}. \ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots(27)$$

On the other hand, we know, from the theorems of matrices, that jj component of the inverse matrix of $\mathbf{X'X}$ is given by

$$(\mathbf{X'X})^{-1}{}_{jj}=(-1)^{j+j}|\mathbf{S}|/|\mathbf{X'X}|. \ \dots\dots\dots\dots\dots\dots\dots\dots\dots(28)$$

Thus from equations (25), (27) and (28), we have

$$\sigma\ \sqrt{(\mathbf{X'X})^{-1}}=K_m V_e/K_{m-1}S_e, \ \dots\dots\dots\dots\dots\dots\dots\dots\dots(29)$$

which, we see in Figure 8, is the hight of the top of the SDE above the hyperplane $\beta_j=0$.

FIG. 8.   RELATIONS AMONG THE jj COMPONENT OF VARIANCE-COVARIANCE
MATRIX, AREA OF THE CROSS-SECTION, AND THE VOLUME OF SDE
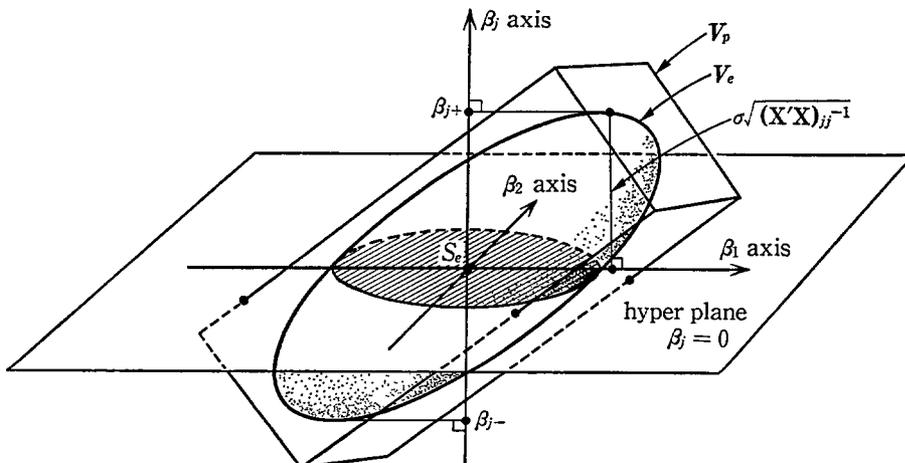
IV. *Application of the SDE Analysis to the Geodetic Observations*

Applying the SDE analysis to actual observations and investigating the shape of the SDE, we can judge the efficiency of observations quantitatively, or we can find an optimum estimation of parameters in a form of their linear combination if the condition of observation is restricted.   In this section we explain how to apply this analysis and then see how useful it is, by using an observation schedule of a geodetic VLBI experiment.

IV.1.   Observational equations of geodetic VLBI

Here the "geodetic VLBI observation" means such an observation as to determine three components of a baseline vector and a clock parameter in the reference framework formed by some radio sources whose coordinates are assumed to be well determined.   Note that the "baseline vector" means a vector connecting two antennas, and the "clock parameter" means the clock difference between the two sites.   By the geodetic VLBI observation we can determine these parameters with an unprecedentedly high accuracy, e.g. with a few centi-meters or with 0.1 nano-second for as long a baseline as intercontinental.

As we know rough values of these parameters, we can choose the unknowns to be their small corrections, thus the observational equations can be linearized.   In the case of our simple model, however, the observational equations are originally linear.   Let us take a reference frame which rotates with a fixed speed with respect to the frame formed by the radio sources.   Choosing the rotation speed as close to that of the Earth as possible, we can treat the baseline vector not changing with time in a relatively short time interval (e.g. a few hours).   Although vectors pointing radio sources are not constant with time, they can be treated as known parameters at a given moment of time, because the rotation speed is fixed.

Let us introduce the observational equations for the geodetic VLBI.   What we observe by the VLBI are the "delay" and the "delay-rate," in which the "delay" means the difference

FIG. 9.   RELATIONS AMONG THE PARAMETERS WHICH APPEAR
IN THE OBSERVATIONAL EQUATIONS
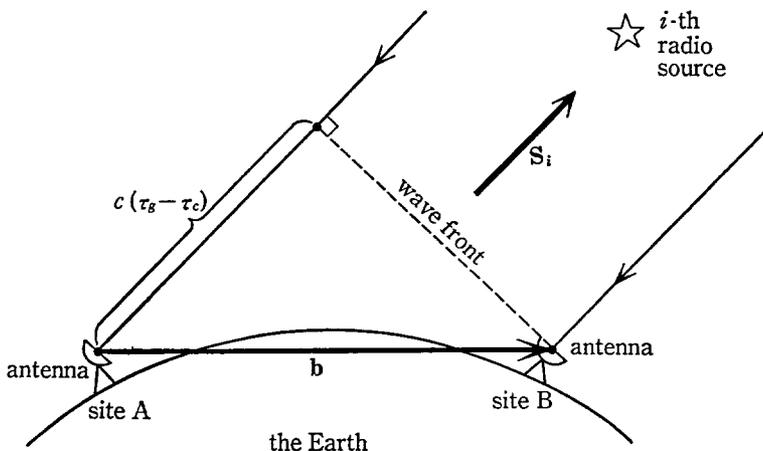
of arrival time of a radio signal between the two antennas.   Here we consider only the case
of delay measurement.   Let s be a unit vector pointing to a radio source, and b be the baseline
vector.   Let τ_g be the observed delay time.   As seen in Figure 9, we have, as an i-th ob-
servational equation,

$$c(\tau_b)_i = \mathbf{b}'\mathbf{s}_i + c\tau_c, \quad \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots(30)$$

where c denotes the light-velocity, and τ_c be the clock advance of the site B against the site
A.   Since the unknown parameters are b and τ_c, we can find the correspondences

$$\left.\begin{array}{l}\boldsymbol{\beta}' = (c\tau_c,\ b_x,\ b_y,\ b_z) \\ \mathbf{X}_i' = (1,\ s_{ix},\ s_{iy},\ s_{iz}) \\ y_i = (c\tau_g)_i\end{array}\right\}, \quad \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots(31)$$

where we take into account the dimensions of parameters.

IV.2.   An example of the VLBI observation schedule used here
    Since we investigate only the SDE (i.e. X'X), we use only the data of the observation
schedule (i.e. a time-table of s_i) and do not use the data of τ_g.   By courtesy of the Radio
Research Laboratory (RRL) of Japan, we can utilize the VLBI observation schedule which
was actually used in an early system-level experiment performed between Kashima branch of
RRL and Mojave station of NASA, USA, on January 23, 1984 (N. Kawano and F. Takahashi
1984, private communication).
    In order to display the distribution of s_i ($i=1,\ \dots,$ 176) in the celestial sphere, let us
define the coordinate system as shown in Figure 10, where the direction of x axis is taken
to be from Kashima to Mojave, z axis toward the center of the common-view zone, and y
axis to form the right-hand system.   Since the distance between the two sites is about 8000km,
and since their latitudes are almost the same, the angle of the common-view zone is about
102° in the x-z plane, and the celestial north pole lies almost in the y-z plane (see Figure 10).
The number of radio sources used in this experiment is 13, and their 176 observed s_i are de-
picted in Figure 11, as the projection onto the x-y plane.
    The whole observations were carried out in 28 hours (about 10 minutes per one observa-
tion), and brought a great success into the pioneering VLBI works of RRL in Japan.

IV.3.   Methods of analysis
    In a general geodetic VLBI experiment all the observations in one or two days' experi-

FIG. 10.   THE COORDINATE SYSTEM AND THE COMMON-VIEW ZONE
FOR KASHIMA-MOJAVE EXPERIMENT
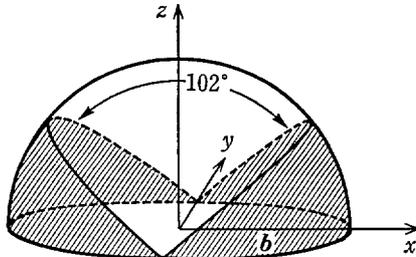
FIG. 11.   THE DISTRIBUTION OF THE POSITIONS OF ALL THE RADIO-SOURCE
OBSERVATIONS IN THE COMMON-VIEW AREA.   THEY ARE PROJECTED
ONTO x-y PLANE
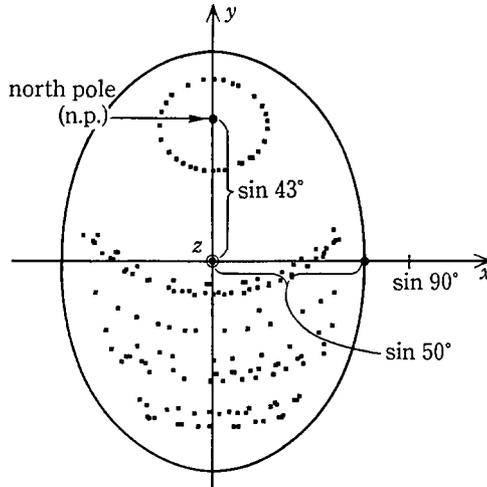
ment are used to make one normal equation of LLST, producing one estimation of **b** in which the theoretical time variations of **b** (such as caused by the Earth tide) are corrected.   In this work, however, we try to divide the experiment into several subsets of observations each of which gives an estimation of **b** and τ_c in a short time-span (e.g. 2 hours).   By such a method of reduction we can see the short-time variation of **b**, and we hope to see, for instance, whether or not an abrupt change is happened just before a great earthquake.   In the former case (i.e. the case of one normal equation) the distribution of s_i is uniform and the error ellipsoid is not elongated so much.   In the latter case (i.e. dividing into many normal equations), however, some s_i's in short time-intervals are not uniform while some are uniform, and then we can see various SDE's and their dependences on s_i distributions.

IV.4.   Results

We divide the whole observations into 14 subsets of 2-hours' observations.   In Figure 12 we show some examples of the s_i distributions.   In Table 1 the eigen values and eigen vectors of the first subset is shown.

It is seen in the Table that two principal axes almost lie in the direction of x and y axes, while the others almost lie in the t-z plane and not directed toward t nor z axes.   As also seen in the Table, one principal axis of the SDE has a very small eigen value, that is, the principal axis is very long (i.e. as long as $\sigma / \sqrt{0.035}$ where σ denotes the expected error of single observation), and another one principal axis is very short (i.e. as short as $\sigma / \sqrt{11.56}$).

Therefore, we depict the SDE in Figure 13 as two cross-sections in the t-z and x-y planes, neglecting those components in eigen vectors which are not arrowed in Table 1.   The length of each principal axis in the Figure is taken to be $\sigma / \sqrt{\lambda_i}$, where λ_i denotes each eigen value.

We also show the SDE of the whole observations in Figure 14.

FIG. 12.   EXAMPLES OF 2-HOURS' SCHEDULES DEPICTED BY THE SAME WAY
AS FIG. 11.   a: THE FIRST 2-HOURS, b: THE SECOND ONE, c: THE WORST
IN THESE FORTEEN 2-HOURS' SCHEDULE, d: THE BEST CASE
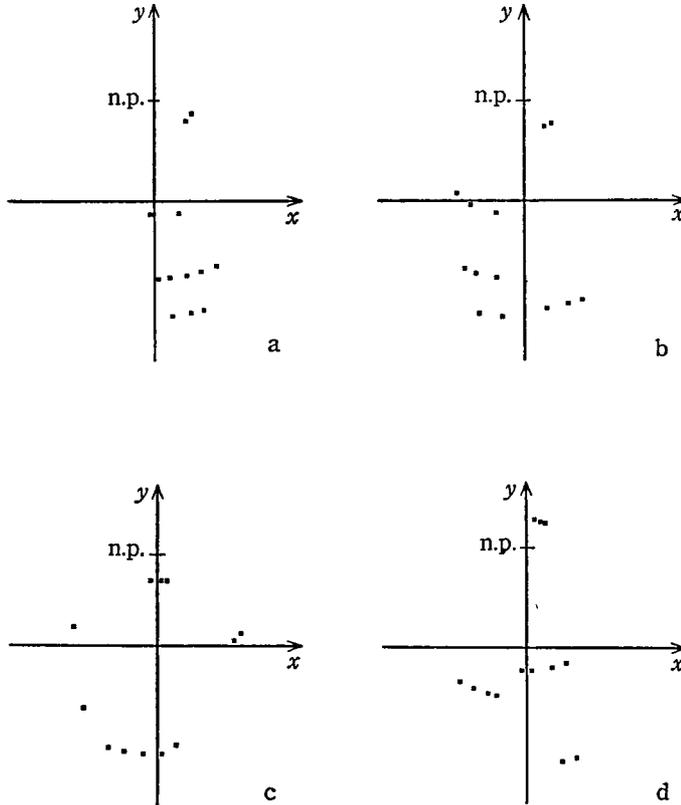
TABLE 1—THE EIGEN VECTORS AND EIGEN VALUES OF AN SDE.
THIS IS CALCULATED WITH THE FIRST 2-HOURS SCHEDULE
WHICH IS SHOWN IN FIGURES 12a AND 13a

EIGEN VECTORS

x:	0.93←	0.04	0.33	0.15
y:	0.04	0.95←	−0.12	−0.29
z:	−0.35	0.28	0.68←	0.58←
t:	0.10	0.14	−0.64←	0.75←

EIGEN VALUES

0.144	1.26	0.035	11.56

IV.5.   Discussions
    Note that the shorter are the lengths of principal axes the better is the composition of

FIG. 13.　CROSS-SECTIONS OF THE SDEs OF OBSERVATION SCHEDULES
SHOWN IN FIG. 12
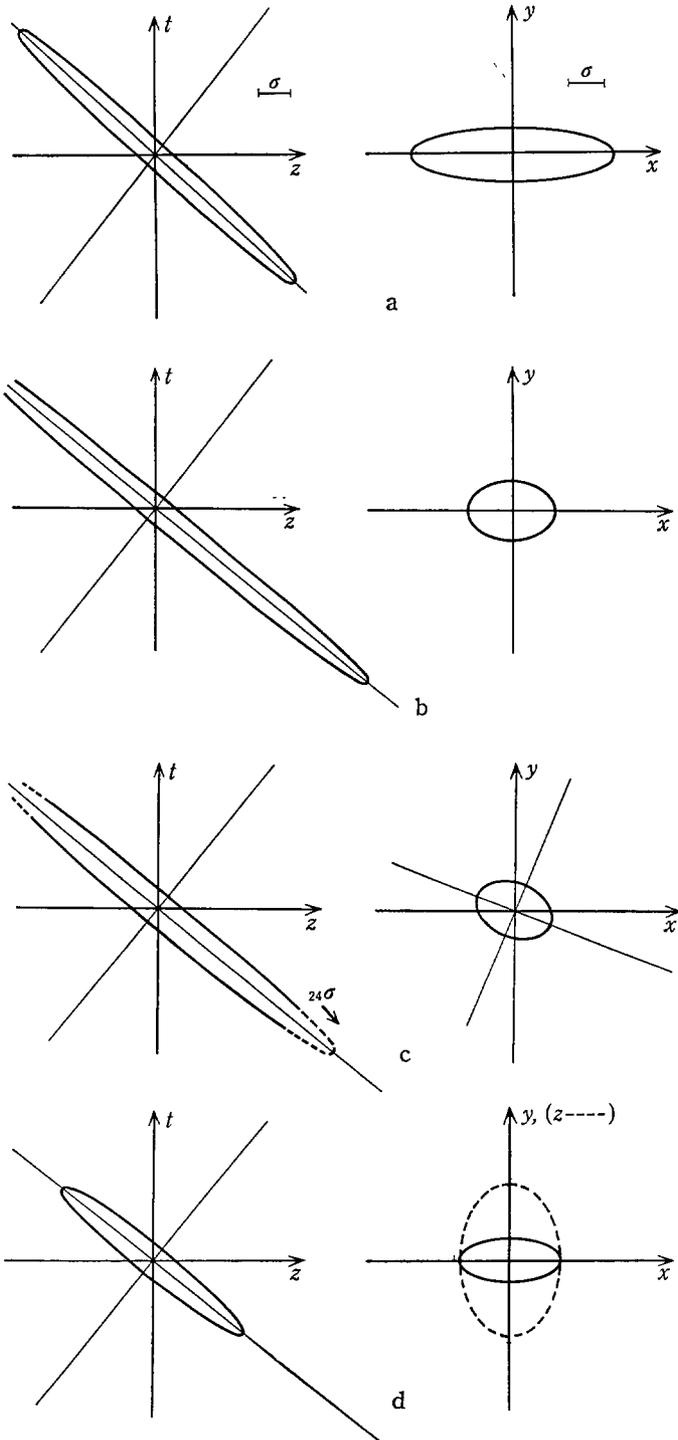
FIG. 14.   THE CROSS-SECTION OF THE SDE OF ALL THE 28-HOURS' OBSERVATIONS
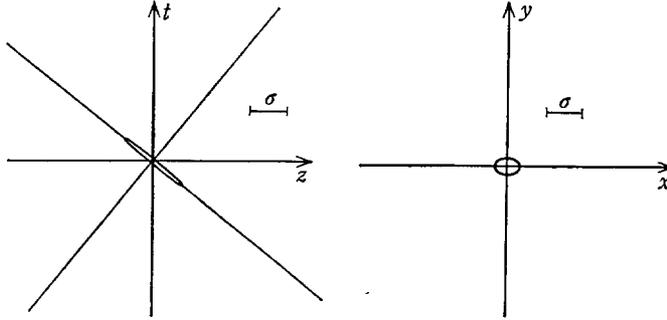
observations.   In Figure 13 we see that the worst case is exemplified by Figure 13c (and Figure 12c), while the best is by Figure 13d (and Figure 12d).   It seems that the reason is because in the former case the distribution of s_i does not fulfill the criterion given by Komaki (1985), i.e. it lacks such s_i as $(0, 0, 1)$, while the latter fulfills the criterion (compare Figure 12c and 12d).   It is noteworthy that the efficiency of the observations can be judged "quantitatively" by the SDE analysis.   However, it should be also noted that the actual criteria for better observations are derived by different investigations such as given by Komaki (1985).

Next, let us show the method to find an optimum estimation of a linear combination of unknown parameters from a given schedule of observations.   As shown in III.3. the s.d.'s of estimated τ_c and b_z are given as the projections of the SDE onto y or z axes, respectively. Therefore they have large errors in the case of Figure 13c.   However, the errors are correlated with each other, that is, if τ_c has a large positive error then b_z has a (absolutely) large negative error.   Thus the error of a combination $c\tau_c + b_z$ becomes rather small.   Such a combination can be find as the direction of the shortest axis of SDE, and its s.d. is calculated from the largest eigen value, λ_{max}, as $\sigma/\sqrt{\lambda_{max}}$.   It is seen in Figure 14 that the geodetic VLBI observations generally provide a higher-accuracy estimation of a combination $c\tau_c + b_z$. If this combination had some geophysical meanings, the VLBI would have been a very sensitive ditector of those geophysical phenomena.

It is noteworthy that if the clocks are synchronized so precisely that τ_c may be treated as a known parameter, the accuracy of the geodetic VLBI can be improved largely, as seen in Figure 14.   This is a reason why the "Connected-element interferometer" at Green Bank (Matsakis et al. 1986) or the "VLA" of the NRAO (Florkowski et al. 1985) in USA produce obsrvations with relatively high accuracy though their baselines are small.

V.   Conclusions

First we explained graphically the method of SDE analysis of the observational equations, in terms of LLST.   We showed relations of some error parameters to the SDE.   It was shown that the s.d. of each unknown parameter is given as the projection of the SDE, then we showed that if the SDE has an elongated axis (which is caused when the combination

of the observational equations is improper), large errors occur in many parameter estimations. The SDE analysis provides a way to judge quantitatively the efficiency of the observational equations.

Looking at the eigen values and eigen vectors calculated in the SDE analysis, we can find an optimum estimation of a linear-combinations of unknown parameters even in the case that the combination of the observational equations is improper.

The SDE analysis was applied to the actual observational equations of a geodetic VLBI experiment, and its usefulness was explained. Some of the characteristics of the geodetic VLBI observations were pointed out.

HITOTSUBASHI UNIVERSITY

ACKNOWLEDGEMENTS

REFERENCES

Bjerhammer, A. 1973, *Theory of Errors and Generalized Matrix Inverses* (Elsevier Scientific Publishing Company, New York).
Florkowski, D.R., Johnston, K.J., Wade, C.M., and de Vegt, C. 1985, *Astron. J.*, **90**, 2381.
Komaki, K. 1985, *J. Geod. Soc. Japan*, **31**, 202.
Matsakis, D.N., Josties, F.J., Angerhofer, P.E., Florkowski, D.R., McCarthy, D.D., Xu Jiayan, and Peng Yunlou 1986, *Astron. J.*, **91**, 1463.
Rao, C.R. 1973, *Linear Statistical Inference and Its Applications,* 2nd edition (John Wiley & Sons, New York).