KONINKLIJK NEDERLANDS METEOROLOGISCH INSTITUUT

MEDEDELINGEN EN VERHANDELINGEN

No. 84

DR. D. J. BOUMAN
DR. H. M. DE JONG

GENERALIZED THEORY OF ADJUSTMENT OF OBSERVATIONS

WITH APPLICATIONS IN METEOROLOGY

1964

F. 11,—

STAATSDRUKKERIJ- EN UITGEVERIJBEDRIJF / 'S-GRAVENHAGE

GENERALIZED THEORY OF ADJUSTMENT OF OBSERVATIONS

WITH APPLICATIONS IN METEOROLOGY

KONINKLIJK NEDERLANDS METEOROLOGISCH INSTITUUT

MEDEDELINGEN EN VERHANDELINGEN

No. 84

DR. D. J. BOUMAN

DR. H. M. DE JONG

GENERALIZED THEORY OF ADJUSTMENT OF OBSERVATIONS

WITH APPLICATIONS IN METEOROLOGY

1964

STAATSDRUKKERIJ- EN UITGEVERIJBEDRIJF / 's-GRAVENHAGE

PUBLIKATIE: K.N.M.I. 102–84

U.D.C.: 519.2:
53.088:
551.5

PREFACE

On occasion of an investigation of the methods of observation in upper air research some rather complex difficulties arose when an analysis was made of the error theory pertaining to these observation methods. After a thorough discussion it was revealed that these difficulties were to be attributed to some fundamental questions concerning the procedure of adjustment of observations which is laid down in many text-books on planning and analysis of experiments. The scheme and rules for adjustment of measured data known to geometers and astronomers and all investigators in science had to be reviewed.

At present the procedure of adjustment is based on the principle of least squares, respectively on an optimal variance reduction using the conditional equations existing between the variables to be measured. It was found that the use of the conditional equations could be removed which has led to the development of a new approach to the problem of adjustment. The results obtained by the new approach appeared to be identical to those of the conventional methods provided that a simple requirement is fulfilled. In this sense the existing theories could be expanded and generalized and their outcome has offered the possibility to find a clear insight in the analysis of all types of measurements ranging from a single observation to the class of repeated indirect conditional observations. The generalized theory throws new light especially on the category of repeated direct and indirect observations which form the majority of observations.

In order to demonstrate the bearing of the generalization the authors have applied the theory to four types of measurements in the field of meteorological practice.

The Director in Chief
Royal Netherlands Meteorological Institute

IR. C. J. WARNERS

CONTENTS

INTRODUCTION

The present study is devoted to some problems which arise in connection with some particular types of measurements in meteorology. In trying to apply the existing theory of adjustment by the method of least squares to these types of measurements the existing theory has been found incomplete insofar that it could not be used to find the answers to some questions of importance. So it seemed worth while trying to generalize the existing theory in order to overcome these difficulties.

The actual problem from which the present study arose is discussed in section 10.3 of Part III (*Applications*). The exact formulation of the questions which could not be answered by using the existing theory is given in Part I (*General Theory*). In this Introduction the main problems discussed in the present study are described in a more informal way, using as an illustration a simple artificial but instructive example.

Let be given that measurements are made on a rectangle $ABCDA$. The measurements give

$$\begin{aligned} AB &= x \\ BC &= y \\ CD &= z \\ DA &= t\,. \end{aligned} \tag{1}$$

For the exact values or "true values" one should have

$$\begin{aligned} \varphi_1 &\equiv x - z = 0 \\ \varphi_2 &\equiv y - t = 0\,. \end{aligned} \tag{2}$$

Suppose one wants to estimate the value of the area A of the rectangle.

$$A = xy \tag{3}$$

This is a problem of adjustment, since the measured values of the length of the sides will, as a rule, on account of the inevitable instrumental and other errors contradict the conditions (2). The method of least squares gives a solution of this problem and the way to obtain the solution has been described many times in readily available text-books *e.g.* Linnik [1961]. However, it seems that this theory cannot answer a series of simple but important questions of which the following are typical specimen.

Suppose one does not use the conditions (2) but the equivalent system

$$\begin{aligned} \varphi_3 &\equiv 3x-5y-3z+5t=0 \\ \varphi_4 &\equiv 2x+3y-2z-3t=0\,, \end{aligned} \tag{4}$$

how does this affect the ultimate solution of the adjustment problem? The system (4) is completely equivalent to the system (2) so one may hope that by using (4) instead of (2) the ultimate solution of the adjustment problem will not be changed but as the present authors could find so far, no proof has been given in literature. It is obvious that the proof should not be confined to the replacement of (2) by (4), but should pertain to a *set* of equivalent systems of conditions out of which (2) and (4) are specific individual systems. This means that prior to the establishment of a proof the set of equivalent systems of conditions ought to be properly defined. This preliminary problem of defining in a proper way the set of equivalent systems of conditions is in itself not easily to solve. For it should be kept in mind that in many practical applications the system of conditions is not given by *linear* equations of the type of the equations (2). It is shown in Part I that the concept of *ideal* in a *ring* is of fundamental importance and that once the concept of ideal is introduced a convenient definition of the set of equivalent systems of conditions may be given and that the solutions obtained by the method of least squares are indeed independent of a specific choice of the system of conditions.

Suppose further that the adjustment by the method of least squares is not based on equation (1) for A and the conditions (2) but on one of the following formulae for A

$$\begin{aligned} A' &= yz \\ A'' &= zt \\ A''' &= xt \\ A^{(\mathrm{iv})} &= \tfrac{1}{2}(xy+zt) \end{aligned} \tag{5}$$

all giving the same numerical value for the exact or true values of x, y, z and t, how does this affect the ultimate solution of the adjustment problem? Each formula of (5) is completely equivalent to formula (3) so one may hope that by using one of the formulae (5) instead of (3) the ultimate solution of the adjustment problem will not be changed, but again no proof seems to exist in literature. Obviously before giving a proof one must have at hand a proper definition of the set of all formulae equivalent in a specific sense to (3). It will be shown, that once a proper definition of the set of conditions equivalent to (2) has been given, it is an easy job to construct a convenient definition of the set of formulae equivalent to (3) and that the ultimate solution of the problem of adjustment is indeed independent of the specific choice of the formula by which the area is computed.

Quite naturally the generalization of the theory of least squares leads to a new

approach of the adjustment problem in which no conditional equations at all are used but a sufficient number of formulae of type (5). It will be proved indeed that the adjustment may be based either on one formula of type (3) and, say, M conditional equations of type (2) or on $(M+1)$ formulae of type (5) provided some quite natural conditions of independence are fulfilled.

In the present study the traditional terminology of "true values", "errors" and "measured values" is used. The authors are fully aware of the fact that these epistemologically somewhat dubious expressions could be eliminated from the theory by using a more sophisticated statistical framework. However, since it is known that these refinements have no influence at all on the final results to be used in applications, it was considered not necessary to burden the first presentation of the generalized theory of adjustment with these details.

As is usually done in studies on the method of least squares, matrices are used throughout. Since no complicated theorems on matrix calculus are needed, a knowledge of the basic principles of matrix calculus is sufficient for the complete understanding of the present work. Furthermore, the reader is expected to have a working knowledge of the ordinary theory of the method of least squares as may be found in text-books as *e.g.* Linnik [1961].

PART I

GENERAL THEORY

1. Equivalent systems of conditional equations

A problem of adjustment always presents itself in the following situation. Measurements have been made on n variables $x_1, \ldots x_n$. For the unknown true values it is a priori known that they are not independent but that M independent relations exist which may be expressed by formulae of the type

$$\varphi_\alpha(x_1, \ldots x_n) = \varphi_\alpha(x_i) = 0\,, \qquad (\alpha = 1, \ldots M)\,. \tag{1.1}$$

One may assume $M < n$, otherwise no measurements are necessary and the values of the n unknowns x_i can be computed from the equations (1.1). Furthermore one may assume that the M relations are independent in that the rank of the matrix $\left| \frac{\partial \varphi_\alpha}{\partial x_i} \right|$, $(\alpha = 1, \ldots M;\ i = 1, \ldots n)$ is equal to M, otherwise one or more relations can be left out of consideration. Given this situation one wants to obtain that value of a given function

$$f = f(x_1, \ldots x_n) \tag{1.2}$$

which may be considered the best value obtainable in a sense which could be specified later on. Instrumental and other errors will cause the measured values $\boldsymbol{x}_1, \ldots \boldsymbol{x}_n$ to differ by an unknown amount from the true values $x_1, \ldots x_n$ so that as a rule (1.1) is violated

$$\varphi_\alpha(\boldsymbol{x}_1, \ldots \boldsymbol{x}_n) \neq 0\,.$$

Whereas the true values $x_1, \ldots x_n$ are pure numbers, the measured values $\boldsymbol{x}_1, \ldots \boldsymbol{x}_n$ are to be considered stochastic variables. It must be emphasized that the stochastic components $\boldsymbol{x}_i - x_i$ are not assumed to be stochastically independent variables; on the contrary, correlation between say $\boldsymbol{x}_i - x_i$ and $\boldsymbol{x}_j - x_j$ is allowed. Furthermore it must be kept in mind that one or more of the conditional equations are also allowed to be of the type

$$\varphi \equiv x_i - x_j = 0\,. \qquad \text{(for specified } i \text{ and } j\text{)} \tag{1.3}$$

Conditional equations of this type express an identity between the true values of certain variables which is a formal way of expressing that some variables have been measured repeatedly. So the theory also covers the case of repeated measurements.

The problems under discussion in the present section pertain to indeterminateness of the system (1.1) of conditional equations. This is not a completely new problem, since an analogous problem is encountered *e.g.* in algebraic geometry. The latter problem being more simple than the former it seems useful first to consider its known solution.

Suppose in the three-dimensional euclidean space of the variables x_1, x_2, x_3 a circle is given by the equations

$$\varphi_1 \equiv x_1^2 + x_2^2 - 1 = 0$$
$$\varphi_2 \equiv x_3 = 0 .$$

The same circle, however, may be given by a multitude of other pairs of equations *e.g.*

$$\varphi'_1 \equiv x_1^2 + x_2^2 + x_3^2 - 1 = 0$$
$$\varphi'_2 \equiv x_3 = 0 ,$$

or

$$\varphi''_1 \equiv x_1^2 + x_2^2 + x_3^2 - x_1 x_3 - 1 = 0$$
$$\varphi''_2 \equiv x_3 = 0 .$$

For the solution of the problem, how to describe the multitude of possibilities of giving the circle by a pair of equations, the algebraic geometers have made use of the theory of polynomial rings and ideals. It is obvious that restriction to polynomial conditional equations would impose limitations on the applicability of the theory of adjustment which are highly undesirable. However, it will be shown that one may use analytic functions as well as polynomials. It further appears that the amount of ideal theory necessary for the solution of the problem is rather small and in fact is restricted to one theorem which will be proved without assuming knowledge of ideal theory with the reader. After these preliminaries the problem will now be tackled formally.

First the concept of a (commutative) *ring* is needed. A *ring* is a set of elements $f, g, h \ldots$ for which two binary operations are defined, called *addition* and *multiplication*, and for which the following axioms are valid:

(*i*) If f and g are elements of the ring, so is $f+g=g+f$ (closure under addition)
(*ii*) $f+(g+h)=(f+g)+h$
(*iii*) There exists a neutral element for addition, written 0, so that $f+0=f$ and $f+(-f)=f-f=0$*)
(*iv*) If f and g are elements of the ring so is $fg=gf$ (closure under multiplication)
(*v*) $f(gh)=(fg)h$
(*vi*) $f(g+h)=fg+fh$
$(f+g)h=fh+gh$.

Examples of rings are the system of (positive and negative) natural numbers,

*) It is not necessary for a ring to have a unit element (neutral element for multiplication). Indeed the ring of multiples of 7 is an example.

the system of natural numbers which are multiples of 7, the system of polynomials in one indeterminate x. The odd natural numbers do not form a ring. A ring may contain subrings (*e.g.* the ring of multiples of 7 is a subring of the ring of all natural numbers). Some subrings are called ideals. A subring I of a ring R is called an ideal if the following condition is fulfilled:
If φ is an element of I than for any f which is an element of R, φf is an element of I.
Examples of ideals are the subring of multiples of 7 in the ring of all natural numbers, the subring of polynomials in one indeterminate x which take the value zero for $x=1$ and $x=2$, in the ring of all polynomials. It is exactly the last example which, properly generalized, makes the ideal theory a useful tool in algebraic geometry. An analogous ideal will solve the problems pertaining to the indeterminateness of the system of conditional equations.
Now consider in the n dimensional space E_n of the variables $x_1, \ldots x_n$ a domain D, which for reasons of simplicity is assumed to be singly connected. A function $F(x_1, \ldots x_n)$ is called analytic on D if to every point $P \equiv (\hat{x}_1, \ldots \hat{x}_n)$ of D there corresponds a neighbourhood of P in which the function can be represented by a convergent power series in $(x_1 - \hat{x}_1, \ldots x_n - \hat{x}_n)$.
It is obvious that the set of all functions analytic on D forms a ring.
Let further M independent functions $\varphi_1, \ldots \varphi_M$ be given analytic on D.

$$\text{By} \qquad \varphi_1 = \varphi_2 \ldots = \varphi_M = 0 \tag{1.4}$$

a $(n-M)$ dimensional variety V is defined in the space E_n consisting of all sets $(x_1, \ldots x_n)$ for which (1.4) is valid. Now the collection of functions analytic on D becoming zero in each point of V forms an ideal. For if ψ and χ are functions becoming zero on V then also $\psi + \chi$ and $\psi\chi$ become zero on V and for each function ψ becoming zero on V and each function F analytic on D also $F\psi$ becomes zero on V. Furthermore the functions $\varphi_1, \ldots \varphi_M$ form a *base* of the ideal as stated by the following theorem.

Theorem of base

If ψ is a function analytic on D becoming zero on V then there exist M analytic functions $l_1, \ldots l_M$ so that

$$\psi = \sum_{\nu=1}^{M} l_\nu \varphi_\nu .$$

Proof. Introduce in E_n a new system of coordinates $(y_1, \ldots y_n)$ by

$$\begin{aligned} y_1 &= \varphi_1(x_1, \ldots x_n) \\ y_2 &= \varphi_2(x_1, \ldots x_n) \\ &\vdots \end{aligned}$$

$$\begin{aligned} y_M &= \varphi_M(x_1, \ldots x_n) \\ y_{M+1} &= x_{M+1} \\ &\vdots \\ y_n &= x_n . \end{aligned} \tag{1.5}$$

In this new system of coordinates the variety V is given by

$$y_1 = y_2 = \ldots y_M = 0 .$$

Let $\psi = \psi(y_1, \ldots y_n)$ be the function ψ expressed in the new coordinates. It is always possible to write ψ in this form since the rank of the matrix $\left| \frac{\partial \varphi_\alpha}{\partial x_i} \right|$, $(\alpha = 1, \ldots M;\ i = 1, \ldots n)$ equals M, from which it follows that the rank of the matrix $\left| \frac{\partial y_j}{\partial x_i} \right|$, $(i, j = 1, \ldots n)$ equals n, so that a unique inverse transformation by analytic functions exists.
Since ψ becomes zero on V one has

$$\psi(0, 0, 0, \ldots 0, y_{M+1}, \ldots y_n) = 0 .$$

Now ψ is analytic on D and must be developable in a power series in $(y_1, \ldots y_n)$. So there exist a development

$$\psi = l_1 y_1 + l_2 y_2 + \ldots l_M y_M \tag{1.6}$$

in which $l_1, \ldots l_M$ are analytic functions
Substituting (1.5) in the development (1.6) one obtains

$$\psi = l_1 \varphi_1 + l_2 \varphi_2 + \ldots l_M \varphi_M$$

which completes the proof.

Remarks:
1. The given proof is an adaptation to the present situation of a proof given by *Kähler* [1934] for an analogous situation.
2. The proof is constructive in that actual determination of the coefficients $l_1, \ldots l_M$ may be done following the procedure indicated in the proof.
By way of illustration let an example follow:

$$\begin{aligned} \varphi_1 &\equiv x_1 - x_2 = 0 \\ \varphi_2 &\equiv x_2 - x_3^2 = 0 \\ \psi &\equiv x_1^2 - x_3^4 = 0 \\ y_1 &= x_1 - x_2 \\ y_2 &= x_2 - x_3^2 \\ y_3 &= x_3 \end{aligned}$$

So

$$x_1 = y_1 + x_2 = y_1 + y_2 + x_3^2 = y_1 + y_2 + y_3^2$$

and

$$\begin{aligned}\psi &= x_1^2 - x_3^4 = (y_1 + y_2 + y_3^2)^2 - y_3^4 = \\ &= (y_1 + y_2)^2 + 2y_3^2(y_1 + y_2) = \\ &= y_1(y_1 + 2y_2 + 2y_3^2) + y_2(y_2 + 2y_3^2) = \\ &= y_1 l_1 + y_2 l_2\end{aligned}$$

with

$$\begin{aligned}l_1 &= y_1 + 2y_2 + 2y_3^2 = x_1 + x_2 \\ l_2 &= y_2 + 2y_3^2 = x_2 + x_3^2\end{aligned}$$

Using the theorem of base the proof of the following theorem may be given which describes the possibilities of selecting another base.

Theorem:

Let Φ be the vector*)

$$\Phi = \underset{M1}{\Phi} = \begin{vmatrix} \varphi_1 \\ \vdots \\ \varphi_M \end{vmatrix}, \tag{1.7}$$

Ψ the vector $\Psi = \underset{M1}{\Psi} = \begin{vmatrix} \psi_1 \\ \vdots \\ \psi_M \end{vmatrix}$,

L the matrix $L = \underset{MM}{L} = \begin{vmatrix} l_{11} \dots l_{1M} \\ \vdots \\ l_{M1} \dots l_{MM} \end{vmatrix}$

of rank M and of which the elements l_{ik} are functions analytic on D. Then the possibilities of choosing a new base are exhausted by the formula

$$\Psi = L\Phi\,. \tag{1.8}$$

Proof.

(1) Suppose Φ is a base. Then since L has the rank M, Ψ is a system of M independent functions becoming zero on $V[\Phi]$ and may therefore be used as base. One simply has to substitute Ψ for Φ in the theorem of base to obtain the

*) Subscripts under the symbol for a matrix indicate the number of rows and columns. As a rule the subscripts are put below the matrix only on its first introduction after which the subscripts are omitted to facilitate reading.

coefficients needed to express an arbitrary function F, zero on V, as a linear function of Ψ.
(2) Suppose Ψ is a base. Then each element of Ψ becomes zero on V, so that the theory of base applies to each element of Ψ in which case (1.8) is only a compact description of the theorem of base.
Finally the following theorem describes the possibility of substituting an equivalent function for the function f (1.2).

Theorem:

If f' is a function equivalent to f in the sense that on V both functions take the same values, then there exist M analytic functions $l_1, \ldots l_M$ so that

$$f' = f + \sum_{\nu=1}^{M} l_\nu \varphi_\nu \,. \tag{1.9}$$

Proof. Apply the theorem of base to $(f' - f)$.

Corollarium: The number of independent functions taking the same values on $V[\Phi]$ equals $(M+1)$.

Summing up the results of the present section it can be stated that the three proved theorems give a complete description of the indeterminateness of both the system of conditional equations (1.1) and of the formula (1.2). In the next sections the question is investigated whether the possibility of using various sets of conditional equations and various formulae will have any consequence for the ultimate results of the procedure of adjustment.

The theory of adjustment of indirect conditioned observations thus may be considered pertaining to problems on the equivalence classes of functions mutually congruent modulo an ideal of functions becoming zero on a given analytic variety. The theorem of base guarantees the existence of a base of the ideal. The theorem on the substitution $f \to f'$ is used to escape explicit introduction of equivalence classes.
One of the main problems in the next sections originates from the circumstance that from

$$f' \equiv f \pmod{I}$$

it does not follow that

$$\frac{\partial f'}{\partial x_i} \equiv \frac{\partial f}{\partial x_i} \pmod{I} .$$

This difficulty will be overcome by introducing additional assumptions which in a natural way arise from statistical requirements (see *e.g.* the argumentation leading to (2.26)).

2. Adjustment and conditional equations

It now becomes necessary to distinguish carefully between the true values $x_1, \ldots x_n$ and the measured values $\boldsymbol{x}_1, \ldots \boldsymbol{x}_n$. This is done by always writing

functions depending on the observed values with explicit indication of the argument $(\boldsymbol{x}_1, \ldots \boldsymbol{x}_n)$. The true values $x_1, \ldots x_n$ form together the vector X

$$X = \underset{1n}{X} = \{x_1, \ldots x_n\} \tag{2.1}$$

and the observed values $\boldsymbol{x}_1, \ldots \boldsymbol{x}_n$ the vector $\boldsymbol{X}$

$$\boldsymbol{X} = \underset{1n}{\boldsymbol{X}} = \{\boldsymbol{x}_1, \ldots \boldsymbol{x}_n\} . \tag{2.2}$$

For the argument $(\boldsymbol{x}_1, \ldots \boldsymbol{x}_n)$ in functions usually $\boldsymbol{X}$ is written in short. Several estimators of the function

$$f = f(x_1, \ldots x_n) = f(X)$$

are considered which are denoted by $\mathscr{S}_1(f)$, $\mathscr{S}_2(f)$ etc.
As a first estimator consider

$$\mathscr{S}_1(f) = f(\boldsymbol{X}) + \Lambda(X)\Phi(\boldsymbol{X}) , \tag{2.3}$$

Φ being the vector already introduced (1.7) and Λ a vector $\Lambda = \underset{1M}{\Lambda}$, to be determined later on.
Denoting the expectation operator by $\mathscr{E}$ one has

$$\mathscr{E}(\boldsymbol{X}) = X .$$

Furthermore it is assumed that

$$\mathscr{E} f(\boldsymbol{X}) = f(X)$$

and

$$\mathscr{E}\Phi(\boldsymbol{X}) = \Phi(X) = 0 . \tag{2.4}$$

These are very restrictive assumptions, but in dealing with non-linear functions f and non-linear conditional equations $\Phi = 0$ they seem to be inevitable. In consequence of the assumptions made one has

$$\mathscr{E}\mathscr{S}_1(f) = \mathscr{E} f(\boldsymbol{X}) + \Lambda(X)\mathscr{E}\Phi(\boldsymbol{X}) = f(X) \tag{2.5}$$

which means that $\mathscr{S}_1(f)$ represents an unbiassed estimator of f.
It is customary in modern statistics to consider in a collection of estimators the best estimator that individual estimator which has the minimum variance. This estimator is called the most efficient estimator of the collection. This in fact is the modern version of the principle of least squares. To select the most efficient estimator out of the collection (2.3) by making a specific choice for the vector Λ it is necessary to obtain a formula for the variance of $\mathscr{S}_1(f)$. The way to obtain this formula is well known. To make the present study readible in itself a complete derivation is given; analogous derivations however are omitted when similar formulae are needed.

By definition

$$\text{var } \mathscr{S}_1(f)=\mathscr{E}\{\mathscr{S}_1(f)-\mathscr{E}\mathscr{S}_1(f)\}^2 . \tag{2.6}$$

Using (2.5) and (2.4) this amounts to

$$\begin{aligned}\text{var } \mathscr{S}_1(f)&=\mathscr{E}\{\mathscr{S}_1(f)-f(X)\}^2= \\ &=\mathscr{E}\{f(\boldsymbol{X})-f(X)+\Lambda(X)\Phi(\boldsymbol{X})\}^2 .\end{aligned} \tag{2.7}$$

It is assumed that $\boldsymbol{X}-X$ is small so that within the order of accuracy needed, one has

$$f(\boldsymbol{X})-f(X)=f_X(\boldsymbol{X}-X)^T \tag{2.8}*)$$

with

$$f_X=\underset{1n}{f_X}=\left|\frac{\partial f}{\partial x_i}\right|, \qquad i=1,\ldots n$$

and

$$\Phi(\boldsymbol{X})=\Phi(X)+\Phi_X(\boldsymbol{X}-X)^T \tag{2.9}$$

with

$$\Phi_X=\underset{Mn}{\Phi_X}=\begin{vmatrix}\dfrac{\partial\varphi_1}{\partial x_1}\cdots\dfrac{\partial\varphi_1}{\partial x_n}\\ \vdots\qquad\quad\vdots\\ \dfrac{\partial\varphi_M}{\partial x_1}\cdots\dfrac{\partial\varphi_M}{\partial x_n}\end{vmatrix} \tag{2.10}$$

Substitution in (2.7) gives

$$\text{var } \mathscr{S}_1(f)=\mathscr{E}\left[f_X(\boldsymbol{X}-X)^T+\Lambda\Phi+\Lambda\Phi_X(\boldsymbol{X}-X)^T\right]^2 .$$

However, since $\Phi(X)=0$, (2.4) this may be simplified to

$$\begin{aligned}\text{var } \mathscr{S}_1(f)&=\mathscr{E}\left[(f_X+\Lambda\Phi_X)(\boldsymbol{X}-X)^T\right]^2= \\ &=\mathscr{E}\left[(f_X+\Lambda\Phi_X)(\boldsymbol{X}-X)^T(\boldsymbol{X}-X)(f_X^T+\Phi_X^T\Lambda^T)\right]= \\ &=\quad(f_X+\Lambda\Phi_X)\mathscr{E}\left[(\boldsymbol{X}-X)^T(\boldsymbol{X}-X)\right]\cdot(f_X^T+\Phi_X^T\Lambda^T) .\end{aligned}$$

The second factor in this product, the factor containing the expectation operator, is exactly what is called by definition the covariance matrix of the variables $x_1,\ldots x_n$, which is denoted here by S.

$$S=\underset{nn}{S}=\mathscr{E}\left[(\boldsymbol{X}-X)^T(\boldsymbol{X}-X)\right]=|\operatorname{cov}(x_i,x_j)| \tag{2.11}$$

so that

$$\text{var } \mathscr{S}_1(f)=(f_X+\Lambda\Phi_X)S(f_X^T+\Phi_X^T\Lambda^T) \tag{2.12}$$

*) The superscript T denotes matrix transposition. As a general rule the suffix X denotes differentiation (gradient) $\dfrac{\partial}{\partial x_i}$.

which is the formula needed for the variance of the estimator $\mathscr{S}_1(f)$. Formulae of the same type will occur in many places in the present study and are in future presented to the reader without proof as immediate consequences of formulae of the type (2.3). From (2.12) it follows that var $\mathscr{S}_1(f)$ is a minimum if

$$\frac{\partial}{\partial \Lambda} \text{ var } \mathscr{S}_1(f) = 2\{\Phi_X S f_X^T + \Phi_X S \Phi_X^T \Lambda^T\} = 0 .$$

Let now a matrix Γ be defined by

$$\Gamma = \underset{MM}{\Gamma} = \Phi_X S \Phi_X^T \tag{2.13}$$

which is nothing but the covariance matrix for $\Phi(X)$.
Then

$$\Phi_X S f_X^T + \Gamma \Lambda^T = 0$$

from which it follows that

$$\Lambda^T = -\Gamma^{-1} \Phi_X S f_X^T \tag{2.14}$$

or

$$\Lambda = -f_X S \Phi_X^T \Gamma^{-1} . \tag{2.15}$$

Here use is made of the fact that both S and Γ are symmetrical matrices so that $\Gamma = \Gamma^T$ and $S = S^T$, facts which are in future used without notification. Substitution of (2.15) in (2.3) gives the final formula for the adjustment

$$\boxed{\mathscr{S}_1(f) = f(X) - f_X S \Phi_X^T \Gamma^{-1} \Phi(X)} \tag{2.16}$$

For future reference also the formula for the variance of $\mathscr{S}_1(f)$ is given. Using (2.12) one obtains

$$\text{var } \mathscr{S}_1(f) = f_X S f_X^T + f_X S \Phi_X^T \Lambda^T + \Lambda \Phi_X S f_X^T + \Lambda \Phi_X S \Phi_X^T \Lambda^T .$$

The first term on the right hand side of this formula represents the variance of $f(X)$ and substitution of (2.14) and (2.15) gives using (2.13)

$$\begin{aligned} \text{var } \mathscr{S}_1(f) &= \text{var } f(X) - f_X S \Phi_X^T \Gamma^{-1} \Phi_X S f_X^T - f_X S \Phi_X^T \Gamma^{-1} \Phi_X S f_X^T + \\ &\quad + f_X S \Phi_X^T \Gamma^{-1} \Phi_X S \Phi_X^T \Gamma^{-1} \Phi_X S f_X^T = \\ &= \text{var } f(X) - 2 f_X S \Phi_X^T \Gamma^{-1} \Phi_X S f_X^T + f_X S \Phi_X^T \Gamma^{-1} \Gamma \Gamma^{-1} \Phi_X S f_X^T \end{aligned}$$

or

$$\boxed{\text{var } \mathscr{S}_1(f) = \text{var } f(X) - f_X S \Phi_X^T \Gamma^{-1} \Phi_X S f_X^T} \tag{2.17}$$

Formulae (2.16) and (2.17) are definitely not new. In fact they are the standard formulae used in the theory of adjustment of indirect conditional observations. It may be noted that it is known that the adjustment results in a real variance reduction as follows from the following arguments.

Γ is a covariance matrix and thus positively definite and so is Γ^{-1}. So for every vector u, $u\Gamma^{-1}u^T \geqq 0$ from which one obtains:

$$\text{var } \mathscr{S}_1(f) \leqq \text{var } f(X) .$$

Before proceeding to discuss whether the choice of a particular basis has any influence on the estimates it is worth while considering two other ways to obtain the formulae (2.16) and (2.17).
By choosing first $f=x_i$ and letting i run from 1 to n one obtains

$$\mathscr{S}_1(X)=X-(S\Phi_X^T\Gamma^{-1}\Phi(X))^T=X-\Phi^T(X)\Gamma^{-1}\Phi_X S \tag{2.18}$$

since under this choice

$$f_X=E$$

E being the unit matrix.
This gives the possibility for constructing a new estimator $\mathscr{S}_2$ for an arbitrary function $f(X)$ by

$$\mathscr{S}_2(f)=f[\mathscr{S}_1(X)]=f\{X-\Phi^T(X)\Gamma^{-1}\Phi_X S\} .$$

Now one may assume that the correction term $-\Phi^T(X)\Gamma^{-1}\Phi_X S$ is small and consider the Taylor development

$$\begin{aligned}\mathscr{S}_2(f) &= f(X)-f_X\{\Phi^T(X)\Gamma^{-1}\Phi_X S\}^T+O\{\Phi^T(X)\Gamma^{-1}\Phi_X S\}^2 \\ &= f(X)-f_X S\Phi_X^T\Gamma^{-1}\Phi(X)+O\{\Phi^T(X)\Gamma^{-1}\Phi_X S\}^2\end{aligned} \tag{2.19}$$

which implies that neglecting higher order terms

$$\mathscr{S}_2(f)=\mathscr{S}_1(f)$$

from which it follows that also

$$\text{var } \mathscr{S}_2(f)=\text{var } \mathscr{S}_1(f) .$$

Another way to obtain (2.16) is by using the method originally developed by Gauss and Lagrange, from whom also the method of least squares received its name. The method consists in trying to construct a correction vector

$$\underset{1n}{\Delta}=\Delta=|\delta_1 \ldots \delta_n| \tag{2.20}$$

to that $\Phi(X+\Delta)=0$.
However, since for non-linear Φ the use of this equation is as a rule prohibitively complicated, this equation is linearized to

$$\Phi(X)+\Phi_X\Delta^T=0 . \tag{2.21}$$

One further introduces a symmetric positive definite matrix G of weights

$$\underset{nn}{G}=G=|G_{ij}| \tag{2.22}$$

and tries to determine the unknown corrections Δ so that

$$\Delta G \Delta^T$$

takes a minimum under the subsidiary conditions (2.21). The minimum problem is solved by using Lagrange's theory of indeterminate multipliers and by considering the function

$$W = \Delta G \Delta^T + K\{\Phi(X) + \Phi_X \Delta^T\}$$

with the elements of the vector $K = \underset{1M}{K} = | \kappa_1, \ldots \kappa_M |$ as indeterminate multipliers.

The condition for minimum W gives

$$\frac{\partial W}{\partial \Delta^T} = 2\Delta G + K\Phi_X = 0 ,$$

so that

$$\Delta = -\tfrac{1}{2} K \Phi_X G^{-1} . \tag{2.23}$$

By substituting this in (2.21) one obtains as the equation determining K

$$\Phi(X) - \tfrac{1}{2}\Phi_X G^{-1} \Phi_X^T K^T = 0$$

from which it follows that

$$K^T = 2(\Phi_X G^{-1} \Phi_X^T)^{-1} \Phi(X)$$

or

$$K = 2\Phi^T(X)(\Phi_X^T G^{-1} \Phi_X)^{-1} .$$

This gives for Δ according to (2.23)

$$\Delta = -\Phi^T(X)(\Phi_X^T G^{-1} \Phi_X)^{-1} \Phi_X G^{-1} .$$

So that one obtains as a new estimator for $f = X$

$$\mathscr{S}_3(X) = X + \Delta = X - \Phi^T(X)(\Phi_X^T G^{-1} \Phi_X)^{-1} \Phi_X G^{-1} . \tag{2.24}$$

This seems to be a completely new estimator. However, if one chooses for G_{ij} as is usually done weights which are proportional to the elements of $S^{-1} = | \operatorname{cov}(x_i, x_j) |^{-1}$ *e.g.*

$$G = \rho^2 S^{-1} ,$$

ρ^2 being an arbitrary proportionality coefficient then $\mathscr{S}_3$ is reduced to

$$\mathscr{S}_3(X) = X - \Phi^T(X)(\Phi_X^T S \Phi_X)^{-1} \Phi_X S .$$

By (2.13) this is equal to

$$\mathscr{S}_3(X) = X - \Phi^T(X) \Gamma^{-1} \Phi_X S$$

which is the same formula as (2.18) so that

$$\mathscr{S}_3(\boldsymbol{X}) = \mathscr{S}_1(\boldsymbol{X}) .$$

Now having seen that all approaches result in the same final formula for the estimation the question of whether the results depend on the specific choice of a particular base becomes of paramount importance. Two theorems are needed to clarify this point.

First invariance theorem

The estimation by $\mathscr{S}_1$, $\mathscr{S}_2$ and $\mathscr{S}_3$ is invariant for change of base.

Proof. Let a change of base be given by a formula of type (1.8)

$$\tilde{\Phi} = L\Phi . \tag{2.25}$$

Now some caution is needed. For according to (2.4) one must have

$$\mathscr{E}\Phi(\boldsymbol{X}) = 0 .$$

The same must be true for the new base

$$\mathscr{E}\tilde{\Phi}(\boldsymbol{X}) = 0 .$$

This may reached by specifying (2.25) to

$$\tilde{\Phi}(\boldsymbol{X}) = L(X)\Phi(\boldsymbol{X}) .$$

Now (2.25) gives

$$\tilde{\Phi}_X = L_X\Phi + L\Phi_X . \tag{2.26}$$

In each given formula in which Φ_X previously occurred it must always be interpreted as $\Phi_X(X)$. This in (2.26) gives

$$\tilde{\Phi}_X(X) = L_X(X)\Phi(X) + L(X)\Phi_X(X) .$$

But $\Phi(X) \equiv 0$ so that

$$\tilde{\Phi}_X = L\Phi_X \tag{2.27}$$

is the formula to be used in studying the effects of the transformation on the estimation. With (2.27) one obtains for the transformation of (2.13)

$$\tilde{\Gamma} = \tilde{\Phi}_X S\tilde{\Phi}_X^T = L\Phi_X S\Phi_X^T L^T = L\Gamma L^T . \tag{2.28}$$

Since L must be chosen so that L^{-1} exists one also has

$$\tilde{\Gamma}^{-1} = (L\Gamma L^T)^{-1} = (L^T)^{-1}\Gamma^{-1}L^{-1} .$$

Finally one obtains for the estimation based on the new base by (2.16)

$$\begin{aligned}\mathscr{S}_1(f) &= f(\boldsymbol{X}) - f_X S\tilde{\Phi}_X^T \tilde{\Gamma}^{-1}\tilde{\Phi}(\boldsymbol{X}) = \\ &= f(\boldsymbol{X}) - f_X S\Phi_X^T L^T (L^T)^{-1}\Gamma^{-1}L^{-1}L\Phi(\boldsymbol{X}) = \\ &= f(\boldsymbol{X}) - f_X S\Phi_X^T \Gamma^{-1}\Phi(\boldsymbol{X}) = \mathscr{S}_1(f)\end{aligned} \tag{2.29}$$

which completes the proof.

Second invariance theorem

The estimation by $\mathscr{S}_1$, $\mathscr{S}_2$ and $\mathscr{S}_3$ is invariant for the substitution of $f(X)$ by an equivalent $\tilde{f}(X)$, the equivalence of f and $\tilde{f}$ being defined by the condition that f and $\tilde{f}$ take the same values on the variety V defined by $\Phi(X) = 0$.

Proof. Let the substitution be given by (see 1.9)

$$\tilde{f} = f + L\Phi\,, \tag{2.30}$$

in which $L = \underset{1M}{L} = \{L_1, \ldots L_M\}$.

According to the assumptions made one must have

$$\mathscr{E}\tilde{f}(\boldsymbol{X}) = \mathscr{E}f(\boldsymbol{X}) = f(X) = \tilde{f}(X)\,.$$

This may be reached by specifying (2.30) to

$$\tilde{f}(\boldsymbol{X}) = f(\boldsymbol{X}) + L(X)\Phi(\boldsymbol{X})\,.$$

Differentiation of (2.30) would give

$$\tilde{f}_X = f_X + L_X\Phi + L\Phi_X\,. \tag{2.31}$$

However, in each given formula in which f_X previously occurred it must always be interpreted as $f_X(X)$. This in (2.31) gives

$$\tilde{f}_X(X) = f_X(X) + L\Phi_X\,.$$

Using this one obtains by (2.29)

$$\begin{aligned}\mathscr{S}_1(\tilde{f}) &= \tilde{f} - \tilde{f}_X S\Phi_X^T \Gamma^{-1}\Phi(\boldsymbol{X}) = \\ &= f + L\Phi(\boldsymbol{X}) - (f_X + L\Phi_X)S\Phi_X^T\Gamma^{-1}\Phi(\boldsymbol{X}) = \\ &= f + L\Phi(\boldsymbol{X}) - f_X S\Phi_X^T\Gamma^{-1}\Phi(\boldsymbol{X}) - L\Phi_X S\Phi_X^T\Gamma^{-1}\Phi(\boldsymbol{X})\,.\end{aligned}$$

Remembering the definition of Γ one observes that in the right hand side the second and the last term are equal except for the sign so that

$$\mathscr{S}_1(\tilde{f}) = \mathscr{S}_1(f)$$

which completes the proof.
For future references two formulae will now be proved.

Define*)

$$\mathscr{S}_X = \left(\frac{\partial \mathscr{S}}{\partial \boldsymbol{X}}\right)_{X=X} = (\text{by } 2.16) = f_X - f_X S\Phi_X^T \Gamma^{-1} \Phi_X \,. \tag{2.32}$$

Then one has

$$\mathscr{S}_X S\Phi_X^T = f_X S\Phi_X^T - f_X S\Phi_X^T \Gamma^{-1} \Phi_X S\Phi_X^T \,,$$

which according to the definition of Γ (2.13) may be simplified to

$$\boxed{\mathscr{S}_X S\Phi_X^T = 0 \,.} \tag{2.33}$$

Furthermore one also has

$$\mathscr{S}_X Sf_X^T = f_X Sf_X^T - f_X S\Phi_X^T \Gamma^{-1} \Phi_X Sf_X^T$$

which by comparison with (2.17) may be identified as

$$\boxed{\mathscr{S}_X Sf_X^T = \text{var } \mathscr{S}(f) \,.} \tag{2.34}$$

Summing up the results of the present section it may be stated that the theorems proved indicate that neither the particular choice of a specific base Φ nor the choice of a specific f has any influence at all on the final result of the adjustment procedure. This means that one has the freedom to choose both for Φ and f those functions which happen to be the most convenient for practical computation. So *e.g.* one may by (2.28) choose a base Φ so that in this base Γ becomes a diagonal matrix, which is of some importance because Γ is throughout the procedure the only matrix of which the inverse is needed.

Finally it must be noted that the development (2.8), the development (2.19) and the linearization (2.21) all serve the same purpose *viz.* to avoid prohibitive complications. If, as is often done, the theory is developed from the onset under the assumptions of linearity of the conditional equations, the real difficulties of the theory are obscured and a clear insight in its limitations becomes impossible.

A good example of the difficulties one encounters even in rather simple problems when linearization is not applied may be found in a paper by Ritsema and Scholte [1961].

It must be observed that if from the onset the conditional equations are given as linear equations the first invariance theorem becomes trivial. For M linear equations define a linear $(n-M)$ dimensional variety L_{n-M}. It is well known that the adjustment procedure then amounts to an orthogonal projection of the point representing the measured values on the linear variety L_{n-M} provided

*) In this formula the index 1, 2 or 3 is omitted, being superfluous.

that orthogonality is defined with respect to a metric given by the inverse S^{-1} of the covariance matrix. The mere fact that the adjustment procedure can be described in terms of the geometry of linear metric spaces proves the invariance theorem.

3. A new approach to the problem of adjustment

Suppose one has at hand a number of N independent formula $f_1, \ldots f_N$ which are equivalent in that they take the same values for the true values. From the correlarium of section 1 it is known that N cannot exceed $(M+1)$. For the present no other relation between N and M will be assumed besides the natural condition

$$N \leqq M+1 .$$

The question arises whether it would be possible to estimate the common value which is assumed by $f_1, \ldots f_N$ for the true values of the independent variables by taking a suitably chosen linear combination of $f_1(X), \ldots f_N(X)$. Define the following notations

$$F = \underset{N1}{F} = \begin{vmatrix} f_1 \\ \vdots \\ f_N \end{vmatrix} \tag{3.1}$$

$$\Lambda = \underset{1N}{\Lambda} = \{\lambda_1, \ldots \lambda_N\} . \tag{3.2}$$

Then the linear combination of $f_1, \ldots f_N$ is given by

$$\mathscr{S}_4(f) = \Lambda F .$$

As usually it is assumed that

$$\mathscr{E} F(X) = F(X) .$$

In order that $\mathscr{S}_4$ be an unbiassed estimator one should have

$$\Lambda = \Lambda(X)$$

and

$$\sum_{i=1}^{N} \lambda_i = 1 . \tag{3.3}$$

The last condition is written as

$$\Lambda e = 1 , \tag{3.4}$$

in which

$$e = \underset{N1}{e} = \begin{vmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{vmatrix} \tag{3.5}$$

So the formula to be used for the estimator $\mathscr{S}_4$ is given by

$$\mathscr{S}_4(f) = \Lambda(X)F(X) = \Lambda F(X) \tag{3.6}$$

with the subsidiary condition (3.4).
As before the weights Λ are determined by the condition that $\mathscr{S}_4$ should have minimum variance. The variance of $\mathscr{S}_4$ is according to well known rules given by

$$\text{var } \mathscr{S}_4(f) = \Lambda F_X S F_X^T \Lambda^T$$

with

$$F_X = F_X(X) = \underset{Nn}{F_X} = \begin{vmatrix} \frac{\partial f_1}{\partial x_1} \cdots \frac{\partial f_1}{\partial x_n} \\ \vdots \quad\quad \vdots \\ \frac{\partial f_N}{\partial x_1} \cdots \frac{\partial f_N}{\partial x_n} \end{vmatrix}$$

Minimizing the variance of $\mathscr{S}_4$ under the subsidiary condition (3.4) is done in the usual way by introducing an indeterminate multiplier ε and minimizing the function W

$$W = \Lambda F_X S F_X^T \Lambda^T + \varepsilon(\Lambda e - 1).$$

For abbreviation the covariance matrix C of F is introduced, defined by

$$C = \underset{NM}{C} = F_X S F_X^T \tag{3.7}$$

so that W is simplified to

$$W = \Lambda C \Lambda^T + \varepsilon(\Lambda e - 1).$$

The condition for minimum now reads

$$\frac{\partial W}{\partial \Lambda} = 0 = 2C\Lambda^T + \varepsilon e = 0$$

with the solution

$$\Lambda^T = -\frac{\varepsilon}{2} C^{-1} e$$

or

$$\Lambda = -\frac{\varepsilon}{2} e^T C^{-1} . \tag{3.8}$$

Substituting this in (3.4) one obtains as the equation determining ε

$$-\frac{\varepsilon}{2} e^T C^{-1} e = 1$$

or

$$\varepsilon = \frac{-2}{e^T C^{-1} e}$$

by which the final form of Λ is found by substitution in (3.8)

$$\Lambda = \frac{e^T C^{-1}}{e^T C^{-1} e}$$

and the final form for $\mathcal{S}_4$ follows from this equation by (3.6)

$$\boxed{\mathcal{S}_4(f) = \frac{e^T C^{-1} F(\boldsymbol{X})}{e^T C^{-1} e}} \tag{3.9}$$

For future reference also the final formula for var $\mathcal{S}_4$ is given.

$$\text{var } \mathcal{S}_4 = \Lambda C \Lambda^T = \frac{e^T C^{-1}}{e^T C^{-1} e} \cdot C \cdot \frac{C^{-1} e}{e^T C^{-1} e} = \frac{e^T C^{-1} e}{(e^T C^{-1} e)^2}$$

or

$$\boxed{\text{var } \mathcal{S}_4 = \frac{1}{e^T C^{-1} e}} \tag{3.10}$$

It might be noted that the denominator of var $\mathcal{S}_4$ *viz.* $e^T C^{-1} e$ is nothing but the sum of all elements of the matrix C^{-1}.

Now again the question arises whether the result is independent of the specific choice of $F = \begin{vmatrix} f_1 \\ \vdots \\ f_N \end{vmatrix}$. By a counter example it is shown that for *arbitrary* N independence may not exist. Take *e.g.* $N=1$, then (3.9) takes the simple form

$$\mathcal{S}_4(f) = f_1(\boldsymbol{X}) ,$$

and clearly by making another choice, say $\tilde{f}_1$, it is possible that

$$\mathcal{S}_4(f) = f_1(\boldsymbol{X}) \neq \tilde{f}_1(\boldsymbol{X}) = \mathcal{S}_4(\tilde{f}) .$$

However, it will be proved in the next section that when making N as large

as possible *i.e.* when choosing $N=M+1$, independence does exist. For the moment the remark must be made that enlarging the number N of functions F results in decreasing (or at least in non-increasing) the invariance of $\mathscr{S}_4$. For increasing N implies that the most efficient estimator determined by minimizing var $\mathscr{S}_4$ is determined out of a larger collection of estimators. So the maximum efficiency is obtained by making $N=M+1$. In case of $N>M+1$, the covariance matrix C which plays a role in (3.9) and (3.10) becomes a singular matrix. This may be proved by remarking that on account of a well-known theorem one has

$$\text{rank } C \leqq \min (\text{rank } F_X,\ \text{rank } S,\ \text{rank } F_X^T) .$$

For $N>M+1$, the rank of F_X and F_X^T is maximally equal to $M+1$, the rank of S is $n \geqq M+1$ from which it follows that the rank of the (N, N) matrix C is maximally equal to $M+1<N$ and consequently $\det(C)=0$. In the next section the relationship between $\mathscr{S}_4$ for $N=M+1$ and $\mathscr{S}_1$ will be investigated and it will be proved that under this condition the new method becomes identical to the original method.

4. Identity of the old and the new approach to the problem of adjustment

As already announced, in this section the identity of the old and the new approach to the problem of adjustment is to be proved, subjected to the condition

$$N=M+1 . \tag{4.1}$$

Two proofs will be given, one based on investigation of the basic assumptions of both methods and one on their final results.

First proof. Consider the basic formula (3.6)

$$\mathscr{S}_4(f)=\Lambda F(\boldsymbol{X})=\lambda_1 f_1(\boldsymbol{X})+\lambda_2 f_2(\boldsymbol{X})+ \ldots \lambda_{M+1} f_{M+1}(\boldsymbol{X}) .$$

Using (3.3) this may be written as

$$\mathscr{S}_4(f)=f_{M+1}(\boldsymbol{X})+\lambda_1\{f_1(\boldsymbol{X})-f_{M+1}(\boldsymbol{X})\}+\lambda_2\{f_2(\boldsymbol{X})-f_{M+1}(\boldsymbol{X})\}+ \ldots + \\ +\lambda_M\{f_M(\boldsymbol{X})-f_{M+1}(\boldsymbol{X})\} .$$

Now the functions

$$f_\nu(\boldsymbol{X})-f_{M+1}(\boldsymbol{X}) \qquad \nu=1, \ldots M$$

are functions becoming zero for the true values X. So according to the theorem of base one may write

$$f_\nu(\boldsymbol{X})-f_{M+1}(\boldsymbol{X})=\sum_{\rho=1}^{M} l_{\nu\rho}\varphi_\rho(\boldsymbol{X}) .$$

Using this $\mathscr{S}_4(f)$ takes the form

$$\mathscr{S}_4(f)=f_{M+1}(\boldsymbol{X})+\sum_{\rho=1}^{M} h_\rho \varphi_\rho .$$

But this is exactly the same form which has been used in the method $\mathscr{S}_1(f)$. So minimizing the variance of $\mathscr{S}_1$ amounts to the same as minimizing the variance of $\mathscr{S}_4$. Remembering that in $\mathscr{S}_1$ it is of no importance which particular choice of f_{M+1} and of φ_ρ is made, it follows that also the results of $\mathscr{S}_4$ are independent of the particular choice of $f_1, \ldots f_{M+1}$.

Second proof. Consider $\mathscr{S}_1(f)$ (see 2.16)).

$$\mathscr{S}_1(f)=f(\boldsymbol{X})-f_X S\Phi_X^T \Gamma^{-1}\Phi(\boldsymbol{X}) . \tag{4.2}$$

It is known from the first invariance theorem that it is of no importance what particular choice of Φ is made. So one is entitled to make the following choice

$$\varphi_\rho=f_\rho-f_{M+1} \qquad (\rho=1, \ldots M) . \tag{4.3}$$

Equation (4.3) may be written as a matrix equation by introducing

$$F=\underset{M+1,1}{F}=\begin{vmatrix} f_1 \\ \vdots \\ f_{M+1} \end{vmatrix} \tag{4.4}$$

and

$$Q=\underset{M,M+1}{Q}=\begin{vmatrix} 1\,0\,0\ldots & 0 & -1 \\ 0\,1\,0 & 0 & -1 \\ 0\,0\,1 & 0 & -1 \\ \cdot & \vdots & \\ \vdots & & \\ \cdot & 0 & -1 \\ 0\,0\,0\ldots & 1 & -1 \end{vmatrix} = \left| E \vdots -e \right| \tag{4.5}$$

viz.

$$\Phi(\boldsymbol{X})=QF(\boldsymbol{X})$$

from which it follows that

$$\Phi_X=QF_X .$$

This gives for Γ, defined by (2.13) using (3.7)

$$\Gamma=\Phi_X S\Phi_X^T=QF_X SF_X^T Q^T=QCQ^T$$

giving for $\mathscr{S}_1(f)$ of (4.2)

$$\mathscr{S}_1(f)=f(\boldsymbol{X})-f_X SF_X^T Q^T(QF_X SF_X^T Q^T)^{-1} QF(\boldsymbol{X}) .$$

In this equation the choice of $f(X)$ is still open. Making $f(X)$ in turn equal to $f_1(X), \ldots f_{M+1}(X)$ one obtains a sequence of formulae which may be given the vectorial form

$$\begin{aligned}\mathscr{S}_1(F) &= F(X) - F_X S F_X^T Q^T (Q F_X S F_X^T Q^T)^{-1} Q F(X) \\ &= F(X) - C Q^T (Q C Q^T)^{-1} Q F(X) .\end{aligned} \tag{4.6}$$

Multiplying this equation to the left by Q one finds

$$\begin{aligned}Q\mathscr{S}_1(F) &= QF(X) - QCQ^T (QCQ^T)^{-1} QF(X) = \\ &= QF(X) - QF(X) = 0 .\end{aligned}$$

Remembering the definition of Q (4.5) one may interpret this as expressing that the vector $\mathscr{S}_1(F)$ consists of identical elements. This is not a surprising result for in fact it is a restatement of the second invariance theorem.
Now (4.6) is giving

$$C^{-1}\mathscr{S}_1(F) = C^{-1}F(X) - Q^T (QCQ^T)^{-1} QF(X)$$

and

$$e^T C^{-1}\mathscr{S}_1(F) = e^T C^{-1} F(X) - e^T Q^T (QCQ^T)^{-1} QF(X) . \tag{4.7}$$

Using the definitions (4.5) and (3.5) one finds that

$$e^T Q^T = (Qe)^T = \begin{vmatrix} 1 & 0 & 0 & \ldots & 0 & -1 \\ 0 & 1 & 0 & & 0 & -1 \\ 0 & 0 & 1 & & 0 & -1 \\ \vdots & & & & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 1 & -1 \end{vmatrix} \begin{vmatrix} 1 \\ 1 \\ 1 \\ \\ 1 \end{vmatrix} = 0$$

so that (4.7) is simplified to

$$e^T C^{-1}\mathscr{S}_1(F) = e^T C^{-1} F(X) . \tag{4.8}$$

It has already been proved that $\mathscr{S}_1(F)$ is a vector consisting of identical elements which means that for $\mathscr{S}_1(F)$ may be written

$$\mathscr{S}_1(F) = e\,\mathscr{S}_1(f_\nu)$$

ν being an arbitrary number out of the series $1, \ldots M+1$.
Substituting this in (4.8) one finally obtains

$$\mathscr{S}_1(f_\nu) = \frac{e^T C^{-1} F(X)}{e^T C^{-1} e} = \mathscr{S}_4(f) \quad \text{for all} \quad \nu = 1, \ldots M+1$$

which completely proves the identity between $\mathscr{S}_1$ and $\mathscr{S}_4$.
The first and the second invariance theorems valid for $\mathscr{S}_1$ now guarantee that also the results of the method $\mathscr{S}_4$ are independent of the specific choice for F.

So the following theorem has been proved:

Third invariance theorem

The estimation by $\mathscr{S}_4$ is invariant for substitution of F by an equivalent $\tilde{F}$, equivalence being defined by the condition that the elements of F and $\tilde{F}$ take the same numerical values for the true values of the independent variables.
Although as an immediate consequence it also follows that var $\mathscr{S}_1 = \text{var } \mathscr{S}_4$ it might be worth while to give also the demonstration by which the formula (2.17) for var $\mathscr{S}_1$

$$\text{var } \mathscr{S}_1 = \text{var } f(\boldsymbol{X}) - f_X S \Phi_X^T \Gamma^{-1} \Phi_X S f_X^T$$

may be transformed into the formula (3.10) for var $\mathscr{S}_4$

$$\text{var } \mathscr{S}_4 = \frac{1}{e^T C^{-1} e}.$$

Using the same substitutions as above one gets for an indeterminate f

$$\begin{aligned} \text{var } \mathscr{S}_1(f) &= \text{var } f(\boldsymbol{X}) - f_X S F_X^T Q^T \Gamma^{-1} Q F_X S f_X^T = \\ &= f_X S f_X^T - f_X S F_X^T Q^T \Gamma^{-1} Q F_X S f_X^T = \\ &= f_X [S - S F_X^T Q^T \Gamma^{-1} Q F_X S] f_X^T . \end{aligned} \tag{4.9}$$

Making f in turn equal to $f_1, \ldots f_M$ the right hand side of (4.9) gives the elements of a matrix which according to the second invariance theorem consists of identical elements. So one may write

$$\begin{aligned} e\, e^T \text{ var } \mathscr{S}_1(f) &= F_X [S - S F_X^T Q^T \Gamma^{-1} Q F_X S] F_X^T = \\ &= F_X S F_X^T - F_X S F_X^T Q^T \Gamma^{-1} Q F_X S F_X^T = \\ &= C - C Q^T \Gamma^{-1} Q C \end{aligned}$$

So

$$e\, e^T C^{-1} e \text{ var } \mathscr{S}_1(f) = e - C Q^T \Gamma^{-1} Q e .$$

Since it has already been demonstrated that $Qe = 0$ this reduces to

$$e\, e^T C^{-1} e \text{ var } \mathscr{S}_1(f) = e . \tag{4.10}$$

So the left hand side of (4.10) is a vector with identical elements, all equal to unity from which it follows immediately that

$$\text{var } \mathscr{S}_1(f) = \frac{1}{e^T C^{-1} e}$$

as has been announced.
For future reference a formula comparable with (2.33) will be derived. Re-

membering the definition given by (2.32) one now may write

$$\mathscr{S}_X = \frac{e^T C^{-1} F_X}{e^T C^{-1} e}.$$

So the application of (2.33) now gives the formula

$$\boxed{e^T C^{-1} F_X S \Phi_X^T = 0\,.} \tag{4.11}$$

The final conclusion of the first part of this study is therefore that there are various approaches to the problem of adjustment of indirect conditional observations, all of which however leading to the same result. A particular problem of adjustment may therefore be transformed by various means into a form suitable for practical computation. For although the various approaches are theoretically equivalent they are far from being equivalent from the computer's point of view. In many cases the three invariance theorems offer possibilities to facilitate computing. Moreover, there is the new approach by which no use is made of the conditional equations and by which only a sequence of equivalent functions is used. The advantages of the new method are that in this method the formulae are symmetric in all functions *f*. This symmetry by which all functions are treated in the same way may be of paramount importance in practical computations. In any case just this symmetry will as a rule simplify the programming of the adjustment procedure for high speed electronic computers. That this is not wishful thinking will become clear when studying the examples given in the last part of the present study.
Finally it must be observed that the new method offers the possibility of reducing the amount of computational work at the expense of loss of efficiency by choosing $N < (M+1)$ which however in general should not be advocated, since the invariance theorems then break down.

PART II

CONDITIONAL AND UNCONDITIONAL OBSERVATIONS

5. Survey

In Part I the problem of adjustment of indirect conditional observations was tackled along different lines of approach and it was shown that entirely different schemes for solution of the problem are leading to the same result. In particular several estimators denoted by $\mathcal{S}_1$, $\mathcal{S}_2$, $\mathcal{S}_3$ and $\mathcal{S}_4$ yield the same minimum variance. The estimators $\mathcal{S}_1$, $\mathcal{S}_2$ and $\mathcal{S}_3$ are well-known standard formulae for adjustment, the estimator $\mathcal{S}_4$ in its most general form appears to be unknown. In any case when occasionally the adjustment procedure $\mathcal{S}_4$ has been used by some author this is done without the knowledge of its equivalence to the more conventional procedure $\mathcal{S}_1$. The theoretical equivalence of $\mathcal{S}_4$ with $\mathcal{S}_1$, $\mathcal{S}_2$ and $\mathcal{S}_3$ could be proved by introducing the concept of ideal in a ring of (analytic) functions and three invariance theorems. The estimator $\mathcal{S}_4$ is discernable from the others by the property that no explicit use is made of the set of conditional equations $\varphi_\alpha=0$ which exist between the true values of the variables to be measured. The only point of contact of $\mathcal{S}_4$ with the set of conditions is the requirement that the total number of equivalent functions should be taken equal to $(M+1)$, where M denotes the *total* number of independent equations. The third invariance theorem guarantees that the adjustment remains unchanged when another set of $(M+1)$ equivalent functions is chosen. Under circumstances *e.g.* when the number of conditional equations is unknown, one can be forced to use a smaller number of equivalent functions at the cost of a smaller reduction of variance and the additional sacrifice of the invariance with respect to a set of equivalent functions. When the number of equivalent functions exceeds $(M+1)$, then the procedure of adjustment fails as the covariance matrix C becomes a singular matrix.

Whereas the third invariance theorem plays a part in $\mathcal{S}_4$, both first and second invariance theorems are effective in regard to the methods $\mathcal{S}_1$, $\mathcal{S}_2$ and $\mathcal{S}_3$. The first theorem states invariance with respect to an equivalent set of conditional equations, the second invariance with respect to an equivalent function.

The special role of the conditional equations $\varphi_\alpha=0$ and their explicit use in the methods $\mathcal{S}_1$, $\mathcal{S}_2$ and $\mathcal{S}_3$ in contrast to a total absence of the conditions

in the method $\mathscr{S}_4$ make it relevant to arrange the methods of adjustment into two categories, viz. the "Φ-approach" and the "f-approach". The former involves the methods $\mathscr{S}_1$, $\mathscr{S}_2$ and $\mathscr{S}_3$, the latter the method $\mathscr{S}_4$.
With this classification in mind the most important results of Part I are:

Φ-approach:

$$\mathscr{S}(f)=f(\boldsymbol{X})-f_X S\Phi_X^T \Gamma^{-1}\Phi(\boldsymbol{X}) \tag{5.1}$$

$$\text{var } \mathscr{S}(f)=\mathscr{S}_X(f)Sf_X^T \tag{5.2}$$

$$\mathscr{S}_X(f)S\Phi_X^T=0 \tag{5.3}$$

f-approach:

$$\mathscr{S}(f)=\frac{e^T C^{-1}F(\boldsymbol{X})}{e^T C^{-1}e} \tag{5.4}$$

$$\text{var } \mathscr{S}(f)=\frac{1}{e^T C^{-1}e} \tag{5.5}$$

$$e^T C^{-1}F_X S\Phi_X^T=0\,. \tag{5.6}$$

It can be remarked that equation (5.3) plays the role of checking the correct answer of $\mathscr{S}(f)$ derived from (5.1). The equation (5.6) involves the set of independent conditions $\Phi=0$ explicitly and as such should not appear in the list of formulae for the f-approach. But when the set of conditions is known in advance the equation can also serve to check the estimator derived from (5.4).
Although theoretically the Φ-approach and f-approach are equivalent it may under circumstances be difficult to decide which of both should be preferred in a particular problem of adjustment. In Part III some arguments are given of one in favour of the other. These arguments may depend on some alternative forms in which the above matrix equations may be converted. When some elementary matrix operations like row- and column operations are performed the matrix equations are changed considerably which make them more suitable for numerical computations. These alternative forms are presented in sections 6 and 7.
It has already been put forward in Part I of the present study that the theory of adjustment of indirect conditional observations also covers those of repeated measurements. It will be shown that by constructing a specific set of conditional equations the well-known category of indirect unconditional observations can be considered a special case of conditional observations. In terms of these special types of observations application of the generalized theory of adjustment leads to similar results as using the conventional method of least squares. In some respect adjustment gives more details than the method of least squares. Section 8 deals with this item and the outline there may be considered a first application of the generalized theory.

6. Alternative forms for the estimator and its variance in the case of explicit use of the conditional equations

One can find alternative forms for $\mathscr{S}(f)$ and var $\mathscr{S}(f)$ in the Φ-approach when some elementary row- and column operations are performed. These new expressions only involve a quotient of two determinants of square matrices, the one in the numerator being a bordered one of the matrix in the denumerator.

For convenience some new notations are added to the existing for Γ and C:

$$\Gamma = \Phi_X S \Phi_X^T \tag{6.1}$$

$$C = F_X S F_X^T \tag{6.2}$$

$$c = f_X S f_X^T \tag{6.3}$$

$$B = \Phi_X S f_X^T \tag{6.4}$$

c is nothing else than the variance of f and consequently appears as a diagonal entry of matrix C. The matrix $B = B_{M1}$ may be interpreted as a "mixed" covariance matrix in terms of f and Φ.

Let Γ_Φ and Γ_B denote the bordered matrices:

$$\Gamma_\Phi = \begin{vmatrix} f(X) & \Phi^T(X) \\ B & \Gamma \end{vmatrix} \tag{6.5}$$

$$\Gamma_B = \begin{vmatrix} c & B^T \\ B & \Gamma \end{vmatrix} \tag{6.6}$$

Then the following theorem will be proved:

$$\boxed{\mathscr{S}(f) = \frac{\det \Gamma_\Phi}{\det \Gamma}} \tag{6.7}$$

$$\boxed{\text{var } \mathscr{S}(f) = \frac{\det \Gamma_B}{\det \Gamma}} \tag{6.8}$$

Consider the matrix Γ_Φ:

$$\Gamma_\Phi = \begin{vmatrix} f(X) & \varphi_1(X) & \varphi_2(X) \dots & \varphi_M(X) \\ B_1 & \gamma_{11} & \gamma_{12} & \gamma_{1M} \\ B_2 & \gamma_{21} & \gamma_{22} & \cdot \\ \vdots & & & \vdots \\ B_M & \gamma_{M1} & \dots & \gamma_{MM} \end{vmatrix} \tag{6.9}$$

To calculate $\det \Gamma_{\Phi}$ first develop the determinant by the elements of the first row, then by the elements of the first column to obtain the coefficients of $f(\boldsymbol{X})$ and $\varphi_{\alpha}(\boldsymbol{X})$:

$$\det \Gamma_{\Phi}=f(\boldsymbol{X}) \det \Gamma-\sum_{i,j=1}^{M} \varphi_j(\boldsymbol{X}) B_i \gamma_{ij}^{*}$$

where γ_{ij}^{*} denotes the *cofactor* of the element γ_{ij} in the matrix Γ. This cofactor is defined in such a way that

$$|\gamma_{ij}^{*}| = \Gamma^{-1} \det \Gamma.$$

With the symmetry of Γ in mind $(\Gamma^{T}=\Gamma)$ one has:

$$\det \Gamma_{\Phi}=f(\boldsymbol{X}) \det \Gamma-\sum_{i,j=1}^{M} \varphi_j(\boldsymbol{X}) B_i (\Gamma^{-1})_{ij} \det \Gamma.$$

In matrix form after substituting for (6.4)

$$\frac{\det \Gamma_{\Phi}}{\det \Gamma} = f(\boldsymbol{X})-f_X S \Phi_X^{T} \Gamma^{-1} \Phi(\boldsymbol{X}).$$

Hence, comparing this result with the expression for $\mathscr{S}(f)$ (5.1) one obtains:

$$\mathscr{S}(f) = \frac{\det \Gamma_{\Phi}}{\det \Gamma}$$

which completes the proof of the first part of the theorem.
Next consider

$$\Gamma_B = \begin{vmatrix} c & B_1 & B_2 & \dots & B_M \\ B_1 & \gamma_{11} & \gamma_{12} & \cdots & \gamma_{1M} \\ B_2 & \gamma_{21} & \gamma_{22} & & \cdot \\ \vdots & & & & \vdots \\ B_M & \gamma_{M1} & \cdots & & \gamma_{MM} \end{vmatrix} \tag{6.10}$$

Again, to calculate $\det \Gamma_B$ first develop the determinant by the elements of the first row, then by the elements of the first column:

$$\begin{aligned} \det \Gamma_B &= c \det \Gamma-\sum_{i,j=1}^{M} B_i B_j \gamma_{ji}^{*} \\ &= c \det \Gamma-\sum_{i,j=1}^{M} B_i B_j (\Gamma^{-1})_{ji} \det \Gamma. \end{aligned}$$

In matrix form, after substituting for (6.3) and (6.4)

$$\frac{\det \Gamma_B}{\det \Gamma} = f_X S f_X^T - f_X S \Phi_X^T \Gamma^{-1} \Phi_X S f_X^T .$$

Comparing this result with (2.17) one has

$$\text{var } \mathscr{S}(f) = \frac{\det \Gamma_B}{\det \Gamma}$$

which completes the proof of the second part of the theorem.

Remark. The above theorem has a lot in common with a well-known theorem due to Neyman-David [1951], concerning the estimation of linear forms in fundamental parameters from indirect observations such as in linear regression theory.

It has in the beginning been assumed that the observations are stochastically dependent. This corresponds to a scheme of stochastically dependent errors with covariance matrix S involving non-diagonal entries. In practice, however, most often the case is encountered, that one deals with stochastically independent observations and consequently with a covariance matrix S which is a diagonal matrix. In Part III all applications to be described are realizations of this system.

In case of independent observations var $\mathscr{S}(f)$ can be presented in still another form which in general is the most suitable for practical computations.

Consider the matrix

$$\Gamma = \Phi_X S \Phi_X^T$$

det Γ is a generalization of the well-known Gram determinant of a system of vectors. If the matrix S happens to be the unit matrix apart from a common factor, det Γ would be a pure Gram determinant for the system of vectors Φ_X and in analytic geometry the determinant could be interpreted as the square of the volume of the M-dimensional parallelepiped with edges Φ_X in n-dimensional Euclidean space E_n. See Birkhoff, Mac Lane [1948]. With respect to the generalized Gram determinant reference is made to a theorem which is characteristic for this type of determinants *cf.* Linnik [1961].

$$\det \Gamma = \sum \sigma_{j_1}^2 \; \sigma_{j_2}^2 \; \ldots \sigma_{j_M}^2 \; \varepsilon_{j_1, \ldots j_M}^2 \tag{6.11}$$

where $\sigma_{j_i}^2$ are diagonal elements of the matrix S:

$$S = \begin{vmatrix} \sigma_1^2 & 0 & \ldots & 0 \\ 0 & \sigma_2^2 & & \vdots \\ \vdots & & \ddots & \\ 0 & \ldots & & \sigma_n^2 \end{vmatrix}$$

and $\varepsilon_{j_1,\ldots j_M}$ represents a minor of order M, obtained from M rows (and M columns) each numbered $j_1\, j_2, \ldots j_M$ in the matrix Φ_X. The summation is over all possible minors of order M, so that the total number of minors amounts to the binomial coefficient

$$\binom{n}{M} = \frac{n!}{M!(n-M)!}$$

Now, if one examines (6.10), the bordered matrix Γ_B can be written as a product of partitioned matrices:

$$\Gamma_B = \left|\begin{array}{c|c} c & B^T \\ \hline B & \Gamma \end{array}\right| = \left|\begin{array}{c|c} f_X S f_X^T & f_X S \Phi_X^T \\ \hline \Phi_X S f_X^T & \Phi_X S \Phi_X^T \end{array}\right| = \left|\begin{array}{c} f_X \\ \hline \Phi_X \end{array}\right| S \left| f_X^T \middle| \Phi_X^T \right| = PSP^T$$

where P denotes the partitioned matrix

$$P = \underset{M+1\ n}{P} = \left|\begin{array}{c} f_X \\ \hline \Phi_X \end{array}\right| \tag{6.12}$$

Obviously $\det \Gamma_B$ is also a generalized Gram determinant for which:

$$\det \Gamma_B = \sum \sigma_{j_1}^2\ \sigma_{j_2}^2 \ldots \sigma_{j_{M+1}}^2\ \varepsilon_{j_1,\ldots j_{M+1}}^{*2}. \tag{6.13}$$

Here $\varepsilon^*_{j_1,\ldots j_{M+1}}$ denotes a minor of order $(M+1)$ consisting of $(M+1)$ rows (and $M+1$) columns in the matrix P. The total number of minors equals the binomial coefficient

$$\binom{n}{M+1} = \frac{n!}{(M+1)!(n-M-1)!}.$$

If S is proportional to the unit matrix the Gram determinant can be interpreted as the square of the $(M+1)$ dimensional volume of the parallelepiped with vertices f_X and Φ_X in n-dimensional Euclidean space E_n.
After these preliminary remarks it is obvious that one may state the following theorem:

$$\boxed{\text{var } \mathscr{S}(f) = \frac{\sum \sigma_{j_1}^2\ \sigma_{j_2}^2 \ldots \sigma_{j_{M+1}}^2\ \varepsilon^*_{j_1,\ldots j_{M+1}}}{\sum \sigma_{j_1}^2\ \sigma_{j_2}^2 \ldots \sigma_{j_M}^2\ \varepsilon^2_{j_1,\ldots j_M}}.} \tag{6.14}$$

In the numerator the sum is taken over $\binom{n}{M+1}$ minors ε^* of order $(M+1)$ in the partitioned matrix P, in the denumerator over the $\binom{n}{M}$ minors ε of order M in the Jacobian Φ_X.
In case that $S=\sigma^2 E$, var $\mathscr{S}(f)$ equals the *altitude* of the $(M+1)$ dimensional parallelepiped, which has the M dimensional parallelepiped with vertices Φ_X as *base* and the component of f_X orthogonal to the base as altitude.

In the estimator $\mathscr{S}(f)$ itself, *cf.* (6.7) only the denumerator is a generalized Gram determinant, the numerator det Γ_{Φ} however is not.
For the sake of completeness it is noted that by applying a suitable transformation a system of stochastically dependent observations may be converted into another system with stochastically independent variables. Then the covariance matrix S is transformed into a diagonal matrix and all formulae and conclusions previously written for the independent observations remain valid. But it may occur that the reduction of work achieved by using the formula (6.14) is not compensated by the extra work to perform the appropriate transformation.

7. Alternative forms for the estimator and its variance in the case of no explicit use of the conditional equations

When adjusting conditional observations the f-approach leads to the same estimation and the same over all reduction in variance as in the Φ-approach, provided that the number of equivalent functions is one more than the number of independent conditions. Let a critical set of functions be:

$$F = \begin{vmatrix} f_1 \\ \vdots \\ f_{M+1} \end{vmatrix}$$

and

$$\hat{F} = \begin{vmatrix} f_1 \\ \vdots \\ f_M \end{vmatrix} \tag{7.1}$$

According to the third invariance theorem the estimator $\mathscr{S}(f)$ and its minimum variance remain unaltered when another set of equivalent functions has been chosen as starting point.
It is to be expected that in the f-approach alternative forms for $\mathscr{S}(f)$ and var $\mathscr{S}(f)$ exist which are analogous to those in the Φ-approach. Before a theorem is presented which corresponds to that in section 6 some new notations must be introduced.
Let Q be the matrix

$$Q = |E\,|-e| = \begin{vmatrix} 1 & 0 & 0 \ldots & 0 & -1 \\ 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 & -1 \\ & \vdots & & & \\ 0 & \ldots & 0 & 1 & -1 \end{vmatrix} \tag{7.2}$$

Then

$$QF = \begin{vmatrix} f_1 - f_{M+1} \\ \vdots \\ f_M - f_{M+1} \end{vmatrix}$$

Put

$$C_{QF} = QCQ^T = QF_X SF_X^T Q^T \tag{7.3}$$

and

$$C_F = \left| \frac{F^T}{QC} \right| . \tag{7.4}$$

The matrix C_{QF} is the covariance matrix of the set of functions QF, *i.e.* the set $f_\rho - f_{M+1}$, $\rho = 1, \ldots M$.

Theorem:

$$\boxed{\mathscr{S}(f) = (-1)^M \frac{\det C_F}{\det C_{QF}}} \tag{7.5}$$

$$\boxed{\text{var } \mathscr{S}(f) = \frac{\det C}{\det C_{QF}}} \tag{7.6}$$

A proof of this theorem could be given within the scheme of the *f*-approach, but an attempt to arrive at a result appears to be laborious. It would be an advantage that such a proof would be independent from the number of functions. Here a proof is given that starts from the formulae (6.7) and (6.8) while two of the invariance theorems are consulted.

According to the first theorem invariance exists with respect to a change of base. As suitable base take

$$\Phi = QF .$$

This is the matrix notation for the set of functions

$$\Phi_\rho = f_\rho - f_{M+1}, \quad \rho = 1, \ldots M .$$

Then in formula (6.5) one has

$$\Gamma = \Phi_X S \Phi_X^T = QF_X SF_X^T Q^T = QCQ^T = C_{QF}$$

respectively

$$B = \Phi_X S f_X^T = QF_X S f_X^T$$

so that

$$\Gamma_\Phi = \left| \begin{array}{c|c} f(X) & \Phi^T(X) \\ \hline B & \Gamma \end{array} \right| = \left| \begin{array}{c|c} f(X) & F(X)^T Q^T \\ \hline QF_X Sf_X^T & QF_X SF_X^T Q^T \end{array} \right| .$$

The second invariance theorem states that the choice of f is immaterial for the problem of adjustment. It may thus be assumed that f is the representative of say f_{M+1}. For the calculation of $\det \Gamma_\Phi$ an elementary column operation is performed consisting of the addition of the first column to each of the next columns. Then with the definition of $\hat{F}$ (7.1) in mind:

$$\det \Gamma_\Phi = \det \left| \begin{array}{c|c} f_{M+1}(X) & \hat{F}(X)^T \\ \hline QF_X Sf_X^T & QF_X S\hat{F}_X^T \end{array} \right| .$$

Next the first column is permuted with the second, the third, etc. until it has become the last column:

$$\det \Gamma_\Phi = (-1)^M \det \left| \begin{array}{c|c} F(X)^T & \\ \hline QF_X SF_X^T & \end{array} \right| = (-1)^M \det C_F .$$

Referring to (6.7) one finally observes that

$$\mathscr{S}(f) = (-1)^M \frac{\det C_F}{\det C_{QF}}$$

completing the proof for the first part of the theorem.
For the second part examine (6.6):

$$\Gamma_B = \left| \begin{array}{cc} c & B^T \\ B & \Gamma \end{array} \right| = \left| \begin{array}{c|c} f_X Sf_X^T & f_X SF_X^T Q^T \\ \hline QF_X Sf_X^T & QF_X SF_X^T Q^T \end{array} \right| .$$

The choice of f is again arbitrary within the class of equivalent functions, so let f represent the function f_{M+1}. When calculating $\det \Gamma_B$ add the first row to each of the next rows:

$$\det \Gamma_B = \det \left| \begin{array}{c|c} f_X Sf_X^T & f_X SF_X^T Q^T \\ \hline \hat{F}_X Sf_X^T & \hat{F}_X SF_X^T Q^T \end{array} \right|$$

where $\hat{F}_X$ is the Jacobian of $\hat{F}$.
Next add the first column to each of the next columns to obtain

$$\det \Gamma_B = \det \left| \begin{array}{c|c} f_{M+1,X} Sf_{M+1,X}^T & f_{M+1,X} S\hat{F}_X^T \\ \hline \hat{F}_X Sf_{M+1,X}^T & \hat{F}_X S\hat{F}_X^T \end{array} \right| . \tag{7.7}$$

When permuting the first column with the second, the third, etc. until the first column appears as the last column and repeating this for the rows, then we observe that on the right hand side of (7.7) the matrix is nothing else than the covariance matrix C for the set of functions F.
Hence

$$\det \Gamma_B = (-1)^{2M} \det C = \det C .$$

So that, referring to (6.8)

$$\text{var } \mathscr{S}(f) = \frac{\det C}{\det C_{QF}}$$

proving the second part of the theorem.
When comparing the results (7.5) and (7.6) with the original expressions (5.4) and (5.5) it is seen, that although the last expressions are very simple, the alternative forms for $\mathscr{S}(f)$ and var $\mathscr{S}(f)$ involve no matrix inversion and addition of matrix elements. However, for the numerical computation of $\mathscr{S}(f)$ with the aid of the alternative formula two determinants must be evaluated whereas in the original formula only one matrix inversion is necessary.
Again the case of independent observations comes to the fore. Then one has to consider a matrix S with only diagonal entries. The numerator and denumerator in formula (7.6) are both generalized Gram determinants. In particular $\det C = \det F_X S F_X^T$ is the generalized Gram determinant of the system of "*vectors*" F_X defined by the rows of the Jacobian F_X. Its value equals:

$$\det C = \sum \sigma_{j_1}^2 \; \sigma_{j_2}^2 \; \ldots \; \sigma_{j_{M+1}}^2 \; \delta_{j_1, \ldots j_{M+1}}^2 . \tag{7.8}$$

The summation must be taken over all minors δ of order $(M+1)$ which are found in the matrix F_X, so that the total number of minors amounts to the binomial coefficient $\binom{n}{M+1}$. In case that $S = \sigma^2 E$, $\det C$ equals the square of the volume of the $(M+1)$ dimensional parallelepiped with vertices $f_{\rho,X}$ $(\rho = 1, \ldots M+1)$. Further the denumerator $\det C_{QF}$ is the generalized Gram determinant of the system of difference vectors QF_X *e.g.* the system $f_{\rho,X} - f_{M+1,X}$, $\rho = 1, \ldots M$. Its value equals:

$$\det C_{QF} = \sum \sigma_{j_1}^2 \; \sigma_{j_2}^2 \; \ldots \; \sigma_{j_M}^2 \; \delta_{j_1, \ldots j_M}^{*2} , \tag{7.9}$$

where the summation extends over all minors δ^* of order M which can be found in the matrix QF_X. The total number of minors in the denumerator amounts to the binomial coefficient $\binom{n}{M}$. Whenever $S = \sigma^2 E$ the value of $\det C_{QF}$ becomes the square of the volume of the M-dimensional parallelepiped spanned by the difference vectors QF_X.
The above formulae (7.8) and (7.9) suggest another form for var $\mathscr{S}(f)$:

$$\boxed{\text{var } \mathscr{S}(f) = \frac{\sum \sigma_{j_1}^2 \, \sigma_{j_2}^2 \ldots \sigma_{j_{M+1}}^2 \, \delta_{j_1, \ldots j_{M+1}}^2}{\sum \sigma_{j_1}^2 \sigma_{j_2}^2 \ldots \sigma_{j_M}^2 \, \delta_{j_1, \ldots j_M}^{*2}}} \tag{7.10}$$

var $\mathscr{S}(f)$ being the quotient of two determinants specified above. It may be noted that when the matrix $S = \sigma^2 E$, then the minimum variance var $\mathscr{S}(f)$ is equal to the square of the component of each of the vectors $f_{\rho,X}$ orthogonal to

the diagonal parallelepiped spanned by the difference vectors QF_X. The second invariance theorem is then geometrically described by the property that all components of the vectors $f_{\rho,X}$ which represent the ascendants of functions pertaining to the class of equivalent functions, orthogonal to the diagonal parallelepiped are equal in length.

The geometrical description for the case that the matrix S is, apart from a multiplicative factor equal to the unit matrix can be expanded somewhat further in order to gain a better understanding of the relationship between the Φ-approach and the f-approach. Whereas in the previous sections attention was paid to the geometrical meaning of var $\mathscr{S}(f)$, the check-formulae (5.3) and (5.6) give some indications of how to fit in the estimator $\mathscr{S}(f)$ itself in the geometrical picture.
The conditional equations $\Phi=0$ determine an $(n-M)$-dimensional variety v in n-dimensional Euclidean space E_n of the independent variables $x_1, x_2, \ldots x_n$. The vectors Φ_X are normal to the variety. Let the function $f(x)$ be represented by a set of scalar surfaces $f(x)=constant$. The true value of f must be found at the $(n-M-1)$ dimensional cross section of an equiscalar surface with the variety v. The ascendant f_X in the point P representing the true value of f is pointed into a direction normal to the equiscalar surface. In the *class* of equivalent functions each function is represented by a set of equiscalar surfaces, but these are mutually connected by the property, that they all have the intersection with the variety v in common, for all $f(x)$ have identical values when true values for the independent variables are substituted. In addition the ascendants $f_{\rho,X}$ are associated to each other by the fact that $\Phi=QF$ is a possible base for the variety v. Furthermore $\Phi_X=QF_X$.
The last relationship means geometrically that the linear space spanned by the system of difference vectors QF_X is the same as that spanned by the vectors Φ_X. In fig. 1 the situation is shown for $M=1$; $n=3$.

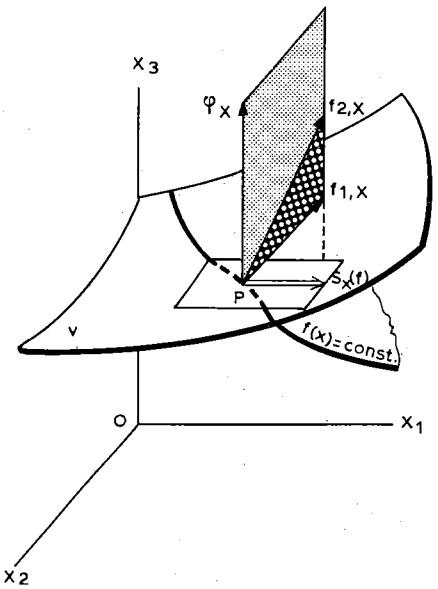

Fig. 1. Geometrical configuration for $M=1$; $n=3$.

Now, according to the checking of formula (5.3) the estimator $\mathscr{S}(f)$ is characterized by the requirement, that its ascendant $\mathscr{S}_X(f)$ is normal to the space spanned by Φ_X, in other words $\mathscr{S}_X(f)$ is incident with the tangential linear space associated with the point P representing the true value on the variety v. $\mathscr{S}(f)$ being an element of the class of equivalent functions, the ascendant $\mathscr{S}_X(f)$ ends in the same linear space as determined by the vectors QF_X. In reference to the above described geometrical interpretation of the Gram determinant the minimum variance is nothing else than the square of the length of the vector $\mathscr{S}_X(f)$. The geometrical interpretation in the Φ-approach and f-approach given for var $\mathscr{S}(f)$ separately, is quite simple, when it is remarked that the altitude of the parallelepiped with vertices f_X and Φ_X, respectively the components of $f_{\rho,X}$ orthogonal to the diagonal parallelepiped with vertices QF_X are equal to the length of $\mathscr{S}_X(f)$.
The characterization of the procedure of adjustment itself is to be found in the way of how the point representing the observed values $x_1, x_2, \ldots x_n$ is displaced with respect to the v-variety. Due to effects of linearization several difficulties arise, especially if the point is not located in the space normal to the Φ-variety.
Probably the best thing to say is that the point is projected onto the tangential linear space associated with the point P representing the true values of $(x_1, x_2, \ldots x_n)$ in the Φ-variety.
In fig. 1, where the Φ-variety is a 2-dimensional surface in E_3, one has to take two functions out of the class of equivalent functions:

$$F = \begin{vmatrix} f_1 \\ f_2 \end{vmatrix} \qquad QF = |f_1 - f_2| .$$

Further

$$F_X = \begin{vmatrix} f_{1,X} \\ f_{2,X} \end{vmatrix}$$

and

$$QF_X = |f_{1,X} - f_{2,X}| .$$

The diagonal parallelogram is here the segment $|f_{1,X} - f_{2,X}|$. This segment is parallel with the ascendant Φ_X. The parallelogram with vertices Φ_X and $f_{1,X}$ has been hatched. The parallelogram (half of it) spanned by the vectors $f_{1,X}$ and $f_{2,X}$ is doubly hatched. The configuration of this figure enables one to identify all formulae in the Φ-approach resp. the f-approach and to study their coherence. For instance formula (5.2), if $S = \sigma^2 E$, is an immediate consequence of the scalar product of two vectors.

8. Unconditional observations

Up to the present the theory of adjustment has been attributed to the category of conditional observations. The adjustment was then partly characterized by the requirement that the corrected values of the variables to be measured or functions of these satisfy the conditions implied by the true values of the variables or at least the linearized forms of these conditional equations.
On the other hand the category of unconditional observations, both direct and indirect exist, which are corrected according to the conventional method of least squares, the theory of which can be found in any textbook on this subject. It may seem at first sight that adjustment for this type of measurements is out of the question but entering into some detail it is observed that these

measurements may also be interpreted as being conditional observations. To this aim the appropriate set of conditions should be detected. It appears that these are of a simple form.

Most generally assume that one wants to obtain the value of a quantity which depends on n variables $x_1, x_2, \ldots x_n$. When the set of variables is measured once, there is no problem to ask for the "best" value of f. Only after the variables have been measured repeatedly the statement of the problem to find the best value of f has any sense. This feature is in contrast with conditional observations where the problem already arises when the variables are measured one time.

In trying to find a solution for the problem of adjustment in this special case both the Φ-approach and f-approach are studied. In view of the rather complex computations the case of a quantity which is a function of one variable is examined first.

One has

$$f=f(x)\,.$$

Let the variable x be measured N times and let N realizations be $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots \boldsymbol{x}_N$. The measured values are considered to be stochastic and from now on each $\boldsymbol{x}_i$ is interpreted as a realization of a *separate* variable x_i with well-defined stochastic properties. Furthermore it is supposed that the measurements are (stochastically) independent and of unequal accuracy, so that the covariance matrix S is a diagonal matrix.

$$S=\begin{vmatrix} \sigma_1^2 & 0 & \ldots & 0 \\ 0 & \sigma_2^2 & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & & & \sigma_N^2 \end{vmatrix}$$

In order to arrive at the set of conditional equations it is observed that an identity must exist between the true values $x_1, x_2, \ldots x_N$.

So one has:

$$\begin{aligned} \varphi_1 &= x_1 - x_N = 0 \\ \varphi_2 &= x_2 - x_N = 0 \\ &\vdots \\ \varphi_{N-1} &= x_{N-1} - x_N = 0\,. \end{aligned} \tag{8.1}$$

The set of conditions $\Phi=0$ consists therefore of $M=N-1$ equations of the form $x_i-x_j=0$. An equivalent set would be:

$$\begin{aligned}\varphi'_1 &= x_2 - x_1 = 0\\ \varphi'_2 &= x_3 - x_1 = 0\\ &\vdots\\ \varphi'_{N-1} &= x_N - x_1 = 0\,.\end{aligned}$$

This set is another base of the ideal $\Phi = 0$ but according to the first invariance theorem this new set will give rise to the same estimator $\mathscr{S}(f)$.
Although a decision respecting the application of either the Φ-approach or the f-approach appears to be in favour of the f-approach both procedures are followed here.
As there are $(N-1)$ conditions the number of equivalent functions to be taken for an optimal adjustment in the f-approach is N. It is relevant to take:

$$F = \begin{vmatrix} f(x_1) \\ f(x_2) \\ \vdots \\ f(x_N) \end{vmatrix}$$

with

$$F(\boldsymbol{x}) = \begin{vmatrix} f(\boldsymbol{x}_1) \\ f(\boldsymbol{x}_2) \\ \vdots \\ f(\boldsymbol{x}_N) \end{vmatrix}$$

and

$$F_X = \begin{vmatrix} \dfrac{\mathrm{d}f}{\mathrm{d}x} & 0 & \dots & 0 \\ 0 & \dfrac{\mathrm{d}f}{\mathrm{d}x} & & \vdots \\ \vdots & & & \\ 0 & \dots & & \dfrac{\mathrm{d}f}{\mathrm{d}x} \end{vmatrix} = \frac{\mathrm{d}f}{\mathrm{d}x} E\,.$$

Then

$$C = F_X S F_X^T = \frac{\mathrm{d}f}{\mathrm{d}x} ESE \frac{\mathrm{d}f}{\mathrm{d}x} = \left(\frac{\mathrm{d}f}{\mathrm{d}x}\right)^2 S \tag{8.2}$$

and

$$C^{-1} = \left(\frac{df}{dx}\right)^{-2} \begin{vmatrix} \frac{1}{\sigma_1^2} & 0 & \dots & 0 \\ 0 & \frac{1}{\sigma_2^2} & & \\ \vdots & & & \\ 0 & & & \frac{1}{\sigma_N^2} \end{vmatrix}$$

For calculating $\mathscr{S}(f)$ and var $\mathscr{S}(f)$ formulae (5.4) and (5.5) are consulted. One has from (8.2)

$$e^T C^{-1} = \left(\frac{df}{dx}\right)^{-2} \left| \frac{1}{\sigma_1^2} \frac{1}{\sigma_2^2} \cdots \frac{1}{\sigma_N^2} \right|$$

and

$$e^T C^{-1} e = \left(\frac{df}{dx}\right)^{-2} \sum_{i=1}^{N} \frac{1}{\sigma_i^2}.$$

Hence

$$\mathscr{S}(f) = \frac{e^T C^{-1}}{e^T C^{-1} e} F(x) = \frac{\sum_{i=1}^{N} \frac{1}{\sigma_i^2} f(x_i)}{\sum_{i=1}^{N} \frac{1}{\sigma_i^2}} \tag{8.3}$$

$$\text{var } \mathscr{S}(f) = \frac{1}{e^T C^{-1} e} = \left(\frac{df}{dx}\right)^2 \frac{1}{\sum_{i=1}^{N} \frac{1}{\sigma_i^2}}. \tag{8.4}$$

Introducing weight factors:

$$g_i = \frac{\sigma^2}{\sigma_i^2} \tag{8.5}$$

σ^2 being the variance for unit weight, the results (8.3) and (8.4) may be written:

$$\mathscr{S}(f) = \frac{\sum_{i=1}^{N} g_i f(x_i)}{\sum_{1}^{N} g_i} = \overline{f(x_i)} \tag{8.6}$$

$$\text{var } \mathscr{S}(f) = \frac{\sigma^2}{\sum_1^N g_i} \left(\frac{\mathrm{d}f}{\mathrm{d}x}\right)^2 . \tag{8.7}$$

Here $\overline{f(x_i)}$ denotes the *weighted mean value* of $f(x_i)$.
These are well-known formulae for estimating according to the method of least squares due to Lagrange-Gauss within the class of indirect observations of one variable of unequal accuracy.

Φ-approach

In the Φ-approach the alternative forms (6.7) and (6.14) are used. As element of the class of equivalent functions $f=f(x_1)$ is taken. The set of conditions is:

$$\Phi = \begin{vmatrix} x_1 - x_N \\ x_2 - x_N \\ \vdots \\ x_{N-1} - x_N \end{vmatrix}$$

and

$$\Phi_X = \begin{vmatrix} 1 & 0 & \cdots & 0 & -1 \\ 0 & 1 & & \vdots & \vdots \\ \vdots & & & & \\ & & & 0 & -1 \\ 0 & & & 0\,1 & -1 \end{vmatrix} = |E| - e| = Q .$$

The matrix Q was already encountered, see (7.2).
The covariance matrix Γ becomes:

$$\Gamma = \Phi_X S \Phi_X^T = QSQ^T = \begin{vmatrix} \sigma_1^2 + \sigma_N^2 & \sigma_N^2 & \cdots & \sigma_N^2 \\ \sigma_N^2 & \sigma_2^2 + \sigma_N^2 & & \sigma_N^2 \\ \vdots & & & \vdots \\ \sigma_N^2 & \sigma_N^2 & & \sigma_{N-1}^2 + \sigma_N^2 \end{vmatrix} \tag{8.8}$$

Further one obtains:

$$P = \underset{NN}{P} = \left| \frac{f_X}{\Phi_X} \right| = \begin{vmatrix} \frac{df}{dx} & 0 \ldots & 0 & 0 \\ 1 & 0 \ldots & 0 & -1 \\ 0 & 1 & \vdots & \vdots \\ \vdots & & 0 & -1 \\ 0 \ldots & & 0\,1 & -1 \end{vmatrix} \tag{8.9}$$

and

$$B = \Phi_X S f_X^T =$$

$$= \begin{vmatrix} 1 & 0 \ldots & 0 & -1 \\ 0 & 1 & 0 & -1 \\ \vdots & & & \vdots \\ & & 0 & -1 \\ 0 \ldots & 0 & 1 & -1 \end{vmatrix} \begin{vmatrix} \sigma_1^2 & 0 & \ldots & 0 \\ 0 & \sigma_2^2 & & \vdots \\ \vdots & & & 0 \\ 0 & \ldots & 0 & \sigma_N^2 \end{vmatrix} \begin{vmatrix} \frac{df}{dx} \\ 0 \\ \vdots \\ 0 \end{vmatrix} = \begin{vmatrix} \sigma_1^2 \frac{df}{dx} \\ 0 \\ \vdots \\ 0 \end{vmatrix}$$

In regard of (6.9)

$$\Gamma_\Phi = \left| \begin{array}{c|c} f(x_1) & \Phi^T(x) \\ \hline B & \Gamma \end{array} \right| =$$

$$= \begin{vmatrix} f(x_1) & x_1 - x_N & x_2 - x_N & \ldots & x_{N-1} - x_N \\ \sigma_1^2 \frac{df}{dx} & \sigma_1^2 + \sigma_N^2 & \sigma_N^2 & & \sigma_N^2 \\ 0 & \sigma_N^2 & \sigma_2^2 + \sigma_N^2 & & \vdots \\ \vdots & \vdots & & \ddots & \\ & & & & \sigma_N^2 \\ 0 & \sigma_N^2 & \ldots & \sigma_N^2 & \sigma_{N-1}^2 + \sigma_N^2 \end{vmatrix} \tag{8.10}$$

Next one has to calculate $\det \Gamma_\Phi$ and $\det \Gamma$.
Γ is a generalized Gram determinant. It is easily computed by application of theorem (6.11).

$$\det \Gamma = \sum \sigma_{j_1}^2 \; \sigma_{j_2}^2 \ldots \sigma_{j_{N-1}}^2 \; \varepsilon_{j_1, \ldots j_{N-1}}^2$$

where ε denotes a minor of order $(N-1)$ in the matrix $\Phi_X = Q = |E| - e|$.

In view of the simple form of Q:

$$\varepsilon^2_{j_1,\dots j_{N-1}}=1\,, \quad \text{for all } j\,,$$

so that

$$\det \Gamma=\sum\sigma^2_{j_1},\sigma^2_{j_2},\dots\sigma^2_{j_{N-1}}=\sum_{j=1}^{N}\ \underset{(j)}{\pi} \tag{8.11}$$

where π is the symbol for the product

$$\pi=\sigma_1^2\ \sigma_2^2\ \dots\sigma_N^2$$

and (j) denotes cancellation in this product of the factor with index j.
To determine $\det \Gamma_\Phi$ develop determinant (8.10) by the elements of the first column

$$\det\Gamma_\Phi=f(x_1)\det\Gamma-\sigma_1^2\frac{\mathrm{d}f}{\mathrm{d}x}\left\{\sum_{j=1}^{N}\ \underset{(1)(j)}{\pi}\ (x_1-x_N)-\ \underset{(1)(2)}{\pi}\ (x_2-x_N)-\dots-\ \underset{(1)(N)}{\pi}\ (x_{N-1}-x_N)\right\}.$$

The coefficients of x_i-x_N are the cofactors of elements in the matrix Γ. For instance referring to (8.8) the coefficient of x_2-x_N becomes

$$(-1)^{2+1}\begin{vmatrix}\sigma_N^2 & \sigma_N^2 & \sigma_N^2 & \dots & \sigma_N^2\\ \sigma_N^2 & \sigma_3^2+\sigma_N^2 & \sigma_N^2 & & \vdots\\ \sigma_N^2 & \sigma_N^2 & \sigma_4^2+\sigma_N^2 & & \\ \vdots & & & \ddots & \\ & & & & \sigma_N^2\\ \sigma_N^2 & \dots & & \sigma_N^2 & \sigma_{N-1}^2+\sigma_N^2\end{vmatrix}=-\begin{vmatrix}\sigma_N^2 & \sigma_N^2 & \sigma_N^2 & \dots & \sigma_N^2\\ 0 & \sigma_3^2 & 0 & & 0\\ 0 & 0 & \sigma_4^2 & & \vdots\\ \vdots & & & \ddots & \\ & & & & 0\\ 0 & \dots & & 0 & \sigma_{N-1}^2\end{vmatrix}=$$

$$=-\sigma_3^2\sigma_4^2\dots\sigma_{N-1}^2\sigma_N^2=-\underset{(1)(2)}{\pi}$$

where $\underset{(1)(2)}{\pi}$ denotes again the product $\sigma_1^2\ \sigma_2^2\dots\sigma_N^2$ but now with cancellation of the factors σ_1^2 and σ_N^2.
Proceeding with the calculation of $\det\Gamma_\Phi$ one finds

$$\det\Gamma_\Phi=f(x_1)\det\Gamma-\frac{\mathrm{d}f}{\mathrm{d}x}\left\{\left(\sum_{2}^{N}\underset{(i)}{\pi}\right)x_1-\sum_{2}^{N}\underset{(i)}{\pi}\,x_i\right\}$$

and applying (6.7) after substituting for (8.11):

$$\mathscr{S}(f)=f(x_1)+\frac{\mathrm{d}f}{\mathrm{d}x}(\bar{x}-x_1)\,. \tag{8.12}$$

Here $\bar{x}$ is the weighted mean value of x_i:

$$\bar{x} = \frac{\sum_{i=1}^{N} \pi_{(i)} x_i}{\sum_{i=1}^{N} \pi_{(i)}} = \frac{\sum_{1}^{N} g_i x_i}{\sum_{1}^{N} g_i}. \tag{8.13}$$

In addition to achieve the minimum variance consider formula (6.14):

$$\text{var } \mathscr{S}(f) = \frac{\sum_{1}^{N} \sigma_{j_1}^2 \ldots \sigma_{j_N}^2 \varepsilon^{*2}_{j_1, \ldots j_N}}{\sum_{1}^{N} \sigma_{j_1}^2 \ldots \sigma_{j_{N-1}}^2 \varepsilon_{j_1, \ldots j_{N-1}}}.$$

Now for ε^* matrix P (8.9) should be examined and for ε the matrix $\Phi_X = Q$. It is easily derived that

$$\text{var } \mathscr{S}(f) = \left(\frac{df}{dx}\right)^2 \frac{\pi}{\sum \pi_{(i)}}$$

or

$$\text{var } \mathscr{S}(f) = \frac{\sigma^2}{\sum g_i} \left(\frac{df}{dx}\right)^2 \tag{8.14}$$

in agreement with (8.7).
When another element out of the class of equivalent functions would have been chosen, *e.g.* $f(x_2)$, then the estimator would have become:

$$\mathscr{S}(f) = f(x_2) + \frac{df}{dx}(\bar{x} - x_2).$$

Similar results are obtained when taking the other function elements. As a special case the function of the weighted mean of x_i can be taken, for $f(\bar{x})$ is also an element of the class of equivalent functions. Substituting for $f(\bar{x})$ gives

$$\mathscr{S}(f) = f(\bar{x}) + \frac{df}{dx}(\bar{x} - \bar{x}).$$

As the second term on the right hand vanishes, one obtains:

$$\mathscr{S}(f) = f(\bar{x}). \tag{8.15}$$

Comparing all the formulae for the estimator it may seem that paradoxically they are not of the same form. In the generalized theory there was an assertion of equality of the estimator $\mathscr{S}(f)$ in the Φ-approach and f-approach. However, it should be borne in mind that the theory guarantees the same estimator so

far linearization is allowed for. As a consequence the estimators give similar results up to higher order terms. When examining formula (8.12) this may be investigated by noting that in practice the numerical value of $\frac{df}{dx}$ is unknown. If then instead of substituting for the true value of the independent variable x in $\frac{df}{dx}$, the measured value x_1 is substituted one obtains the approximation

$$\mathscr{S}(f)=f(x_1)+\left(\frac{df}{dx}\right)_{x_1}(\bar{x}-x_1).$$

Then the right hand side is nothing else than an expansion of $f(x)$ into a Taylor series up to the second order term with respect to the realization x_1. The same applies when the adjustment is based on x_2, x_3 etc.
Writing successively:

$$\begin{aligned}
\mathscr{S}(f)&=f(x_1)+\frac{df}{dx}(\bar{x}-x_1)\\
\mathscr{S}(f)&=f(x_2)+\frac{df}{dx}(\bar{x}-x_2)\\
&\vdots\\
\mathscr{S}(f)&=f(x_N)+\frac{df}{dx}(\bar{x}-x_N).
\end{aligned}\tag{8.16}$$

Multiplying the first equation with g_1, the second with g_2, ... the last with g_N and adding, one finds:

$$\mathscr{S}(f)\sum_i g_i=\sum_i g_i f(x_i)+\frac{df}{dx}\left(\sum_i g_i\bar{x}-\sum_i g_i x_i\right)$$

or, dividing by $\sum_i g_i$:

$$\mathscr{S}(f)=\frac{\sum_i g_i f(x_i)}{\sum_i g_i}=\overline{f(x_i)}$$

This results in the estimator (8.6) of the f-approach. It must be stressed however that in (8.16) second order terms are neglected, so that $f(\bar{x})$ and $\overline{f(x)}$ are estimators of the problem in so far as deletion of terms of the second order and higher is allowed for.
The category of "*unconditional*" observations involves amongst others the direct observations, whether with equal accuracy or not. It turns out that then both the Φ-approach and f-approach give exactly the same outcome.

Indeed,

$$f(x)=x\,.$$

Hence

$$\mathscr{S}(f)=\bar{f}=\frac{\sum_i g_i x_i}{\sum_i g_i}$$

or

$$\mathscr{S}(f)=\bar{x} \tag{8.17}$$

$\bar{x}$ is again the weighted mean of x_i.
The minimum variance is according to (8.14):

$$\text{var } \mathscr{S}(f)=\frac{\sigma^2}{\sum_i g_i} \tag{8.18}$$

so that the estimator $\mathscr{S}(f)$ is the weighted mean of x_i and the "weight" of the weighted mean is equal to the sum of the weights of observation.
In case of equal accuracy the well-known results are obtained:

$$\mathscr{S}(f)=\bar{x} \tag{8.19}$$

$$\text{var } \mathscr{S}(f)=\frac{\sigma^2}{N} \tag{8.20}$$

with

$$\bar{x}=\frac{\Sigma x_i}{N}\,.$$

It has been assumed throughout this investigation that the observations were stochastically independent, respectively that an appropriate transformation had been applied which changed the covariance matrix S into a diagonal matrix. Such a transformation may be avoided in advance and a more direct solution of the problem pursued. It is not to be expected that the formulae for $\mathscr{S}(f)$ and var $\mathscr{S}(f)$ may then be presented in such simple forms as (8.6) and (8.7). Indeed, when the Φ-approach is considered it is difficult to find formulae like (8.12) and (8.15). When however use is made of the f-approach the result still becomes intelligible.
Indeed, referring to (8.2) one has

$$C=\frac{\mathrm{d}f}{\mathrm{d}x}ESE\frac{\mathrm{d}f}{\mathrm{d}x}=\left(\frac{\mathrm{d}f}{\mathrm{d}x}\right)^2 S\,.$$

Hence

$$C^{-1} = \left(\frac{df}{dx}\right)^{-2} S^{-1}$$

$$\mathscr{S}(f) = \frac{e^T S^{-1} F(\boldsymbol{x})}{e^T S^{-1} e} \tag{8.21}$$

$$\text{var } \mathscr{S}(f) = \frac{1}{e^T S^{-1} e} \left(\frac{df}{dx}\right)^2 . \tag{8.22}$$

For direct measurements, where

$$\boldsymbol{x} = \begin{vmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \\ \vdots \\ \boldsymbol{x}_N \end{vmatrix}$$

and

$$\frac{df}{dx} = 1 ,$$

$$\mathscr{S}(x) = \frac{e^T S^{-1} \boldsymbol{x}}{e^T S^{-1} e} \tag{8.23}$$

$$\text{var } \mathscr{S}(x) = \frac{1}{e^T S^{-1} e} . \tag{8.24}$$

In words: in direct, stochastically dependent observations a linear combination of observed values gives rise to a minimum variance, if the coefficients in the linear form are partial sums of column elements with respect to the sum of all elements of the inverse of the covariance matrix S. The reciprocal value of the sum of all elements of the inverse covariance matrix S represents the minimum variance itself.

When proceeding with the more general case of "*unconditional*" observations, in particular the direct observations of more than one variable of unequal accuracy, then the theory of adjustment becomes rather complicated. Most remarkable the f-approach then seems to be shunted by the Φ-approach.

Let a quantity f be given which depends on n variables $x_1, x_2, \ldots x_n$:

$$f = f(x_1, x_2, \ldots x_n) .$$

Further assume that the variables $x_1, \ldots x_n$ are measured N times. Consequently one has at one's disposal the group of observations

$$\begin{array}{l} \boldsymbol{x}_{11} \ldots \boldsymbol{x}_{1n} \\ \boldsymbol{x}_{21} \ldots \boldsymbol{x}_{2n} \\ \vdots \\ \boldsymbol{x}_{N1} \ldots \boldsymbol{x}_{Nn} \end{array}$$

In the Φ-approach each $\boldsymbol{x}_{ij}$ is considered a realization of a separate (stochastic) variable x_{ij}, $i=1, \ldots N$, $j=1, \ldots n$. Moreover the covariance matrix S is supposed to be a diagonal matrix, the measurement of some x_{ij} is therefore not influenced by another variable. When not dealing with stochastically independent observations a suitable transformation can be found which changes the matrix S into a diagonal matrix. Then the adjustment is performed for the new set of variables which are introduced by the transformation.
Between the variables x_{ij} conditions exist expressing the fact that the true values x_{ij} for fixed j are equal:

$$\begin{array}{llll} \varphi_{11} = x_{11} - x_{N1} = 0 & \ldots & \varphi_{1n} = x_{1n} - x_{Nn} = 0 \\ \varphi_{21} = x_{21} - x_{N1} = 0 & & \varphi_{2n} = x_{2n} - x_{Nn} = 0 \\ \vdots & & \vdots \\ \varphi_{N-1,1} = x_{N-1,1} - x_{N1} = 0 & & \varphi_{N-1,n} = x_{N-1,n} - x_{Nn} = 0 \,. \end{array}$$

The total number of these conditional equations amounts to

$$M = n(N-1) \,.$$

Should one proceed in the f-approach then the number of equivalent functions is:

$$M+1 = nN - n + 1 \tag{8.25}$$

A collection of N functions *e.g.* the collection $f(\boldsymbol{x}_{11}, \ldots \boldsymbol{x}_{1n}), f(\boldsymbol{x}_{21}, \ldots \boldsymbol{x}_{2n}) \ldots f(\boldsymbol{x}_{N1}, \ldots \boldsymbol{x}_{Nn})$ is therefore insufficient, it should be enlarged with functions like $f(\boldsymbol{x}_{21}, \boldsymbol{x}_{12}, \ldots \boldsymbol{x}_{1n}), f(\boldsymbol{x}_{31}, \boldsymbol{x}_{12}, \ldots \boldsymbol{x}_{1n}) \ldots f(\boldsymbol{x}_{N1}, \boldsymbol{x}_{12}, \ldots \boldsymbol{x}_{1n})$ and so on, until the total number has increased to $nN-n+1$.
For example, consider the measurement of the area of a rectangle $ABCDA$, described in the introduction. Let $AB=x$ and $BC=y$ be measured twice with unequal accuracy and let the realizations be $\boldsymbol{x}_1, \boldsymbol{y}_1$ resp. $\boldsymbol{x}_2, \boldsymbol{y}_2$. Then the value of the area A of the rectangle is

$$\begin{array}{l} A_1 = \boldsymbol{x}_1 \boldsymbol{y}_1 \\ A_2 = \boldsymbol{x}_2 \boldsymbol{y}_2 \\ A_3 = \boldsymbol{x}_1 \boldsymbol{y}_2 \\ A_4 = \boldsymbol{x}_2 \boldsymbol{y}_1 \\ A_5 = \frac{1}{2}(\boldsymbol{x}_1 + \boldsymbol{x}_2)\boldsymbol{y}_1 \\ \ldots\ldots \end{array}$$

The conditional equations are:

$$\varphi_1 = x_2 - x_1 = 0 \qquad \varphi_2 = y_2 - y_1 = 0$$

So that $nN - n + 1 = 2 \times 2 - 2 + 1 = 3$.

Consequently the "best" linear combination of A_i in the f-approach must be calculated for *three* representants, say A_1, A_2 and A_3. If one should have taken two values, *e.g.* A_1 and A_2, then the reduction of the variance had not been optimal. More than three values leads to a singular covariance matrix C.

Due to the rather queer number of equivalent functions the simplicity of the f-approach is disturbed. The Jacobian F_X involves non-diagonal entries and the calculation of the inverse matrix of C appears to be a difficult task. On the other hand the Φ-approach still leads to an intelligible result.

Let the nN variables be arranged in the series

$$x_{11}, \ldots x_{N1}, \quad x_{12}, \ldots x_{N2}, \ldots x_{1n}, \ldots x_{Nn}$$

and let $f(x_{11}, \ldots, x_{1n})$ be an element of the class of equivalent functions. Then one has

$$\Phi = \underset{n(N-1),1}{\Phi} = \begin{vmatrix} x_{11} - x_{N1} \\ \vdots \\ x_{N-1,1} - x_{N1} \\ \cdots\cdots\cdots\cdots \\ x_{12} - x_{N2} \\ \vdots \\ x_{N-1,2} - x_{N2} \\ \cdots\cdots\cdots\cdots \\ \vdots \\ \cdots\cdots\cdots\cdots \\ x_{1n} - x_{Nn} \\ \vdots \\ x_{N-1,n} - x_{Nn} \end{vmatrix}$$

$$\Phi_X = \begin{vmatrix} \boxed{Q} & 0 \ldots & 0 \\ 0 & \boxed{Q} & \vdots \\ \vdots & & 0 \\ 0 \ldots & & \boxed{Q} \end{vmatrix}$$

where Q denotes as usual the matrix $|E| - e|$ defined in (7.2).
The error matrix is

$$S = \underset{nN,nN}{S} = \begin{vmatrix} S_1 & 0 & \dots & 0 \\ 0 & S_2 & & \\ \vdots & & \ddots & \\ 0 & \dots & & S_n \end{vmatrix}$$

in which for $k = 1, 2 \dots n$

$$S_k = \begin{vmatrix} \sigma_{1k}^2 & 0 & \dots & 0 \\ 0 & \sigma_{2k}^2 & & \vdots \\ \vdots & & \ddots & \\ 0 & \dots & & \sigma_{Nk}^2 \end{vmatrix}$$

To determine $\mathscr{S}(f)$ one should have at one's disposal the covariance matrix Γ and matrix Γ_Φ.

$$\Gamma = \Phi_X S \Phi_X^T = \begin{vmatrix} \boxed{QS_1Q^T} & 0 \dots & & 0 \\ 0 & \boxed{QS_2Q^T} & & \\ \vdots & & & \\ & & & 0 \\ & & \boxed{QS_nQ^T} & \\ 0 \dots & & 0 & \end{vmatrix}$$

To obtain Γ_Φ, first consider the "mixed" covariance matrix:

$$B=\Phi_X S f_X^T = \begin{vmatrix} Q & 0\ldots & 0 \\ 0 & Q & \vdots \\ \vdots & & 0 \\ 0\ldots & 0 & Q \end{vmatrix} \begin{vmatrix} S_1 & 0\ldots & 0 \\ 0 & S_2 & \vdots \\ \vdots & & 0 \\ 0 & 0 & S_n \end{vmatrix} \begin{vmatrix} \frac{\partial f}{\partial x_1} \\ 0 \\ \vdots \\ 0 \\ \frac{\partial f}{\partial x_2} \\ 0 \\ \vdots \\ \frac{\partial f}{\partial x_n} \\ 0 \\ \vdots \\ 0 \end{vmatrix} = \begin{vmatrix} \sigma_{11}^2 \frac{\partial f}{\partial x_1} \\ 0 \\ \vdots \\ 0 \\ \sigma_{12}^2 \frac{\partial f}{\partial x_2} \\ 0 \\ \vdots \\ \sigma_{1n}^2 \frac{\partial f}{\partial x_n} \\ 0 \\ \vdots \\ 0 \end{vmatrix}$$

Then referring to (6.8) matrix Γ_Φ becomes:

$$\Gamma_\Phi = \underset{n(N-1)+1,\,n(N-1)+1}{\Gamma_\Phi} = \begin{vmatrix} f(x_{11},\ldots x_{1n}) & x_{11}-x_{N1}\ldots & & x_{N-1,n}-x_{Nn} \\ \sigma_{11}^2 \frac{\partial f}{\partial x_1} & & 0\ldots & 0 \\ 0 & QS_1Q^T & & \vdots \\ \vdots & & & \\ 0 & & & \\ \sigma_{12}^2 \frac{\partial f}{\partial x_2} & 0 & & \\ 0 & \vdots & QS_2Q^T & \\ \vdots & & & \\ 0 & & & \\ \ldots\ldots & & & \\ \vdots & & & \\ \ldots\ldots & & & \\ \sigma_{1n}^2 \frac{\partial f}{\partial x_n} & & & 0 \\ 0 & & & \\ \vdots & & & QS_nQ^T \\ 0 & 0\ldots & & 0 \end{vmatrix}$$

Finally, using $\mathscr{S}(f) = \dfrac{\det \Gamma_\Phi}{\det \Gamma}$ one should calculate both $\det \Gamma_\Phi$ and $\det \Gamma$. Developing Γ_Φ by the elements of the first column leads to an expression similar to (8.12) but expanded for n variables. The result is

$$\mathscr{S}(f)=f(x_{11}, \ldots x_{1n})+\sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\bar{x}_i-x_{1i}) \tag{8.26}$$

where $\bar{x}_i$ is the weighted mean value of the group of realizations for x_i:

$$\bar{x}_i=\frac{\sum_{k=1}^{N} g_{ki} x_{ki}}{\sum_{k=1}^{N} g_{ki}}, \qquad i=1, \ldots n$$

$$g_{ki}=\frac{\sigma^2}{\sigma_{ki}^2}$$

The partial derivatives for true values being unknown, one puts for x_i the measured value x_{1i} to obtain the approximation:

$$\mathscr{S}(f)=f(x_{11}, \ldots x_{1n})+\sum_{i=1}^{n} \left(\frac{\partial f}{\partial x_i}\right)_{x_{1i}} (\bar{x}_i-x_{1i}) .$$

On the right hand side one discerns the development of $f(\bar{x}_1, \bar{x}_2, \ldots \bar{x}_n)$ into a Taylor series up to the second order term.
However, since $f(\bar{x}_1, \bar{x}_2, \ldots \bar{x}_n)$ is also an equivalent function the estimator for $f(\bar{x}_1, \bar{x}_2, \ldots \bar{x}_2)$ becomes:

$$\mathscr{S}(f)=f(\bar{x}_1, \ldots \bar{x}_n)+\sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\bar{x}_i-\bar{x}_i) .$$

The second part on the right vanishes, proving that $f(\bar{x}_1, \bar{x}_2, \ldots \bar{x}_n)$ itself is an estimator for indirect observations of unequal accuracy.
Again it is found that although theoretically the estimator should be independent of the elements of the class of equivalent functions, the expressions for $\mathscr{S}(f)$ are only comparable inasmuch an expansion into a Taylor series is allowed for and terms of order two and higher may be neglected. Whereas in practice in the expressions for $\mathscr{S}(f)$ the true values should be replaced by measured values, this is superfluous for $\mathscr{S}(f)=f(\bar{x}_1, \bar{x}_2, \ldots \bar{x}_n)$.
The minimum variance var $\mathscr{S}(f)$ will this time be calculated by application of formula (5.2).
One has (cf. 8.26):

$$\mathscr{S}_X(f) = \left| \frac{g_{11}}{\sum_j g_{j1}} \frac{\partial f}{\partial x_1} \cdots \frac{g_{N1}}{\sum_j g_{j1}} \frac{\partial f}{\partial x_1} \frac{g_{12}}{\sum_j g_{j2}} \frac{\partial f}{\partial x_2} \cdots \frac{g_{N2}}{\sum_j g_{j2}} \frac{\partial f}{\partial x_2} \cdots \right.$$
$$\left. \cdots \frac{g_{1n}}{\sum_j g_{jn}} \frac{\partial f}{\partial x_n} \cdots \frac{g_{Nn}}{\sum_j g_{jn}} \frac{\partial f}{\partial x_n} \right|,$$

$$f_X = \left| \frac{\partial f}{\partial x_1} 0 \ldots 0 \quad \frac{\partial f}{\partial x_2} 0 \ldots 0 \ldots \frac{\partial f}{\partial x_n} 0 \ldots 0 \right|$$

$$S = \begin{vmatrix} \boxed{\begin{matrix} \frac{\sigma^2}{g_{11}} & & 0 \\ & \ddots & \\ 0 & & \frac{\sigma^2}{g_{N1}} \end{matrix}} & 0 \ldots & 0 \\ \vdots & & \vdots \\ & & 0 \\ 0 & \boxed{\begin{matrix} \frac{\sigma^2}{g_{1n}} & & 0 \\ & \ddots & \\ 0 & & \frac{\sigma^2}{g_{Nn}} \end{matrix}} & \\ \vdots & & \\ 0 \ldots & & \end{vmatrix}$$

So that from

$$\text{var } \mathscr{S}(f) = \mathscr{S}_X(f) S f_X^T ,$$

one obtains

$$\text{var } \mathscr{S}(f) = \frac{\sigma^2}{\sum_j g_{j1}} \left(\frac{\partial f}{\partial x_1} \right)^2 + \frac{\sigma^2}{\sum_j g_{j2}} \left(\frac{\partial f}{\partial x_2} \right)^2 + \cdots \frac{\sigma^2}{\sum_j g_{jn}} \left(\frac{\partial f}{\partial x_n} \right)^2 ,$$

or

$$\text{var } \mathscr{S}(f) = \sigma^2 \sum_{i=1}^{n} \frac{1}{\sum_{j=1}^{N} g_{ji}} \left(\frac{\partial f}{\partial x_i} \right)^2 \tag{8.27}$$

showing that var $\mathscr{S}(f)$ is to be considered here a generalization of the well-known law of propagation of errors. Indeed, when one has to deal with

indirect observations of equal accuracy for the variables separately:

$$g_{ji}=g_i\,,\quad \text{all } j$$

the variance becomes:

$$\text{var } \mathscr{S}(f)=\frac{\sigma^2}{N}\sum_{i=1}^{n}\frac{1}{g_i}\left(\frac{\partial f}{\partial x_i}\right)^2. \tag{8.28}$$

In this section it has been demonstrated that after admitting conditional equations of the form $\varphi=x_i-x_j=0$ the category of direct and indirect "unconditional" observations is to be interpreted as merely a special case of the wide class of conditional observations. Knowing this, a new aspect of the problem of adjustment presents itself, when the class of conditional observations is enlarged with repeated measurements. Then the set $\Phi=0$ associated with the conditional observations is extended with equations of the above-mentioned type $x_i-x_j=0$ and the Jacobian Φ_X involves amongst others the sub-matrices $Q=|E|-e|$. At the same time the number of equivalent functions in the f-approach is enlarged. This feature suggests a uniform solution of all types of measurements occurring in practice by the method of adjustment and it will be possible to make a working scheme of a uniform method of adjustment for computation by means of high speed electronic computers.

PART III

APPLICATIONS TO SOME METEOROLOGICAL OBSERVATIONS

9. General

The method of adjustment of observations has found wide application in the analyses and planning of experiments. With the appearance of high speed electronic computers the evaluation of quantitative data has got a strong impetus and at present procedures of adjustment are routine pursuits for astronomers, surveyors and all investigators in natural and technical science. The methods of adjustment have not penetrated far into the field of meteorological observations, at least when the measurements are of a complex nature. It is superfluous to say that it is not our aim to review here all types of measurement in the light of the theories developed in the previous sections but by way of demonstration a few of these are investigated in more detail.

Pure conditional observations with non-trivial conditional equations are more exception than the rule. The relationships existing between the base variables can sometimes be rather camouflaged and have to be digged up by a thorough analysis of the problem. It may occur that the number of conditions is large and their precise form unknown. On the other hand a quantity to be determined, which is a function of the base variables, may show alternatives which give rise to the same final result. A nice example is found in modern radar radiosonde observations where the height of the balloon target may be derived from the radar data *c.q.* the temperature, pressure, humidity recordings. When such alternative forms for a quantity are discovered one immediately knows that one deals with the case of conditional observations. The conditions may have their origin from different sources, for instance from a coordinate transformation, a geometrical structure, a physical law, a statistical requirement. When abstaining from a set of conditions the procedure of adjustment can be handled deliberately by using a collection of equivalent functions. This is connected with the choice between the Φ-approach and f-approach.

When for a particular problem both the Φ-approach and f-approach are applied it is not to be expected that the final results, qua form, will be identical. This statement seems at first sight to be somewhat paradoxical in view of the theoretical equivalence of both methods. But it should be kept in mind that all estimations underlying the equivalent methods of adjustment are equal so far as their expansions into Taylor series correspond to each other up to the

the second order term. This is the reason why an estimator in the f-approach may be of a simpler form than an estimator in the Φ-approach and vice versa. It is even possible that estimators in the f-approach differ from each other as well as estimators in the Φ-approach due to the fact that the final result depends on the element of the class of functions on which the calculations are based resp. on the equivalent set of conditional equations. Moreover, since the estimator involves terms which are functions of the true values of the variables, one can alter these — and by consequence the estimator itself — by substituting for a conditional relation between the variables. Then the transformed estimator is equal to the original one modulo Φ.

Further the final expressions for $\mathscr{S}(f)$ involving terms with true values, in practice have to be approximated by some substitution of measured data. This substitution is not uniquely defined.

Owing to these shortcomings, which are caused by linearization processes, it is necessary to know in advance whether the precision of the observation of the base variables is not too low. Otherwise it is recommended to improve the characteristics of the measuring devices considerably.

Referring to the "*residual error*" var $\mathscr{S}(f)$ it should be pointed out that the interpretation of any quantitative result depends on the modern methods of probability theory and mathematical statistics. The authors have consciously abstained from a description of these aspects.

The decision as to whether the Φ-approach or the f-approach is to be preferred depends particularly on the overall scheme and rules of the adjustment procedure. Whereas in general the conditional equations are of a simpler structure than the elements of the class of equivalent functions, the matrix calculus in the f-approach may win in simplicity from the corresponding computation scheme in the Φ-approach. In the f-approach the order of the covariance matrix Γ is one unit higher than the order of the covariance matrix C. On the other hand the Jacobian F_X may contain several zero entries, for one may find elements in the class of functions which depend on a reduced number of base variables, which is not always the case with the conditions $\Phi=0$. Another important argument for one method in favour of the other is the final form of the estimator. In the f-approach this form shows much symmetry. In other cases, when several quantities being each functions of the same base variables are to be adjusted simultaneously, one may prefer to adjust the base variables proper, after which the estimated values of the quantities are obtained by merely substituting for the corrected values of the variables in the functional relations, which determine the appropriate quantities (method $\mathscr{S}_2$). When the conditional equations are unknown, then, of course, the only possibility is to use the f-approach, with the risk however that the *number* of equivalent functions is not optimal. A method to determine the optimal number of functions with some certainty is to investigate the rank of matrix Γ with

increasing number of equivalent functions, for this rank is maximally equal to the total number of conditions plus one.
Before proceeding with some applications in the field of meteorological measurements the scheme and rules of both the equivalent adjustment procedures are presented in matrix terminology.

Summary of Formulae and Rules of Adjustment in Matrix Terminology

Φ-approach

Variables: $x_1, \dots x_n$
Covariance matrix: S
Quantity to be determined: $f(x_1, \dots x_n)$

Find conditional equations: $\Phi = \begin{vmatrix} \varphi_1 \\ \vdots \\ \varphi_M \end{vmatrix} = 0$

Determine Jacobian: Φ_X
Calculate covariance matrix: $\Gamma = \Phi_X S \Phi_X^T$

Two alternatives

a

Calculate: f_X
Determine mixed covariance matrix: $f_X S \Phi_X^T$
Find: Γ^{-1}
Calculate: $f_X S \Phi_X \Gamma^{-1}$
Next: $f_X S \Phi_X \Gamma^{-1} \Phi(x)$
To obtain: $\mathscr{S}(f) = f(x) - f_X S \Phi_X \Gamma^{-1} \Phi(x)$
and var $\mathscr{S}(f) = \mathscr{S}_X S f_X^T$
Check: $\mathscr{S}_X S \Phi_X^T = 0$

b

Calculate: $B = \Phi_X S f_X^T$
and $c = f_X S f_X^T$

Put: $\Gamma_\Phi = \begin{vmatrix} f(x) & \Phi^T(x) \\ B & \Gamma \end{vmatrix}$

and: $\Gamma_B = \begin{vmatrix} c & B^T \\ B & \Gamma \end{vmatrix}$

to obtain: $\mathscr{S}(f) = \dfrac{\det \Gamma_\Phi}{\det \Gamma}$

and var $\mathscr{S}(f) = \dfrac{\det \Gamma_\Phi}{\det \Gamma}$

f-approach

Variables: $x_1, \dots x_n$
Covariance matrix: S
Quantity to be determined: $f(x_1, \dots x_n)$

Determine equivalent functions: $F = \begin{vmatrix} f_1 \\ \vdots \\ f_{M+1} \end{vmatrix}$

Determine Jacobian: F_X
Calculate covariance matrix: $C = F_X S F_X^T$

Two alternatives

a

Calculate: C^{-1}
Determine: $e^T C^{-1}$
and $e^T C^{-1} e$

to obtain: $\mathscr{S}(f) = \dfrac{e^T C^{-1} F(x)}{e^T C^{-1} e}$

and var $\mathscr{S}(f) = \dfrac{1}{e^T C^{-1} e}$

Check: $e^T C^{-1} F_X S \Phi_X^T = 0$

b

Find: QF_X
Calculate: $QC = QF_X S F_X^T$
Put: $C_{QF} = QF_X S F_X^T Q^T$

and $C_F = \begin{vmatrix} F^T \\ \cdots \\ QC \end{vmatrix}$

to obtain: $\mathscr{S}(f) = (-1)^M \dfrac{\det C_F}{\det C_{QF}}$

and var $\mathscr{S}(f) = \dfrac{\det C}{\det C_{QF}}$

Special case: S diagonal

$$\text{var } \mathscr{S}(f) = \frac{\sum_j \sigma^2_{j_1} \cdots \sigma^2_{j_{M+1}} \varepsilon^{*2}_{j_1,\ldots j_{M+1}}}{\sum_j \sigma^2_{j_1} \cdots \sigma^2_{j_M} \varepsilon^2_{j_1,\ldots j_{M+1}}}$$

with ε^*, $(M+1, M+1)$ minor in $P = \begin{vmatrix} f_X \\ \cdots \\ \Phi_X \end{vmatrix}$

and ε, (M, M) minor in Φ_X

Special case: S diagonal

$$\text{var } \mathscr{S}(f) = \frac{\sum_j \sigma^2_{j_1} \cdots \sigma^2_{j_{M+1}} \delta^{*2}_{j_1,\ldots j_{M+1}}}{\sum_j \sigma^2_{j_1} \cdots \sigma^2_{j_M} \delta^2_{j_1,\ldots j_M}}$$

with δ^*, $(M+1, M+1)$ minor in F_X
and δ, (M, M) matrix in QF_X

10. Examples of application in meteorological measurements

10.1. *Upper wind measurement with Radar and Radiosonde*

The conventional method of upper wind evaluation involves the determination of the projection of the sounding- or pilot balloon on a horizontal surface. Azimuth α, slant range r and elevation ε are measured by radar detection and recorded at the end of each time interval (usually one minute). These measured data suffice to compute the balloon's trajectory, but when in particular a sounding balloon is tracked one has an additional quantity at one's disposal, notably the balloon height. This height is obtained from the pressure, temperature and humidity recordings using the well-known hydrostatic equation. The number of data needed for the evaluation of the upper wind has become supernumerary and the system pertains to the class of conditional observations. A mere inspection of the problem of adjustment teaches that the adjustment will be independent of the azimuth and that r, ε and h constitute the set of base variables. Neglecting the earth's curvature the position of the balloon is given by the (cylinder) coordinates (α, d, h), where d represents the horizontal distance of the balloon with respect to the radar site. The horizontal projection of the trajectory is the curve (d, α) in polar coordinates. After the horizontal projection of the trajectory has been determined, the wind is derived from successive positions on this track. It is also necessary to indicate the height for which the wind applies so that the adjustment procedure should be examined for both the distance d and height h.

The true values of r, ε and h are mutually connected by the conditional equation:

$$\varphi = h - r \sin \varepsilon = 0 .$$

Further, considering the rectangle triangle OBB' in fig. 2, one obtains

$$d_1 = r \cos \varepsilon$$
$$d_2 = h \cot \varepsilon$$
$$d_3 = (r^2 - h^2)^{\frac{1}{2}}$$

In the terminology used in the present study d_1 and d_2 represent equivalent

functions in the f-approach. Both the Φ-approach and f-approach can be applied, the f-approach will be preferred as to obtain a result which shows a linear combination of the measured distances. According to theory the total number of equivalent functions should be *two*, as there exists only one condition ($M=1$). Here d_1 and d_2 are taken. It is assumed that the variables are

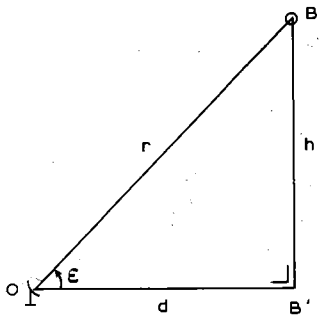

Fig. 2. Balloon's position with respect to the radar site.

not (stochastically) dependent, that the measurement of *e.g.* the azimuth is not dependent on the measurement of the elevation. The covariance matrix S is therefore a diagonal matrix.
Referring to the scheme at the end of section 9 one has:

Variables:

$$x_1=r \quad x_2=\varepsilon \quad x_3=h\,.$$

Covariance matrix:

$$S = \begin{vmatrix} \sigma_r^2 & 0 & 0 \\ 0 & \sigma_\varepsilon^2 & 0 \\ 0 & 0 & \sigma_h^2 \end{vmatrix}$$

Set of equivalent functions:

$$F = \begin{vmatrix} r\cos\varepsilon \\ h\cot\varepsilon \end{vmatrix}$$

Jacobian:

$$F_X = \begin{vmatrix} \cos\varepsilon & -r\sin\varepsilon & 0 \\ 0 & -r\,\mathrm{cosec}\,\varepsilon & \cot\varepsilon \end{vmatrix}$$

Covariance matrix:

$$C = F_X S F_X^T = \begin{vmatrix} \cos^2\varepsilon\sigma_r^2 + r^2\sin^2\varepsilon\sigma_\varepsilon^2 & r^2\sigma_\varepsilon^2 \\ r^2\sigma_\varepsilon^2 & r^2\operatorname{cosec}^2\varepsilon\sigma_\varepsilon^2 + \cot^2\varepsilon\sigma_h^2 \end{vmatrix}.$$

Further, with

$$Q = |E| - e| = |1, -1|$$

$$QC = |\cos^2\varepsilon(\sigma_r^2 - r^2\sigma_\varepsilon^2),\ -\cot^2\varepsilon(\sigma_h^2 + r^2\sigma_\varepsilon^2)|$$

and

$$C_{QF} = QCQ^T = \cot^2\varepsilon(\sin^2\varepsilon\sigma_r^2 + r^2\cos^2\varepsilon\sigma_\varepsilon^2 + \sigma_h^2)$$

Further

$$C_F = \begin{vmatrix} \boldsymbol{d}_1 & \boldsymbol{d}_2 \\ \cos^2\varepsilon(\sigma_r^2 - r^2\sigma_\varepsilon^2) & -\cot^2\varepsilon(\sigma_h^2 + r^2\sigma_\varepsilon^2) \end{vmatrix}$$

C is a (generalized) Gram determinant

$$\det C = \sum_j \sigma_{j_1}^2\sigma_{j_2}^2\delta_{j_1 j_2}^2,$$

where $\delta_{j_1 j_2}$ is a (2, 2) minor in F_X.
Hence one obtains:

$$\det C = \cot^2\varepsilon(r^2\sigma_\varepsilon^2\sigma_r^2 + r^2\sin^2\varepsilon\sigma_\varepsilon^2\sigma_h^2 + \cos^2\varepsilon\sigma_r^2\sigma_h^2).$$

And finally, since $M = 1$,

$$\mathscr{S}(d) = \frac{-\det C_F}{\det C_{QF}},$$

or:

$$\mathscr{S}(d) = \frac{r^2\sigma_\varepsilon^2 + \sigma_h^2}{\sin^2\varepsilon\sigma_r^2 + r^2\cos^2\varepsilon\sigma_\varepsilon^2 + \sigma_h^2}\boldsymbol{d}_1 + \frac{(\sigma_r^2 - r^2\sigma_\varepsilon^2)\sin^2\varepsilon}{\sin^2\varepsilon\sigma_r^2 + r^2\cos^2\varepsilon\sigma_r^2 + \sigma_h^2}\boldsymbol{d}_2. \tag{10.1}$$

The minimum variance becomes:

$$\operatorname{var}\mathscr{S}(d) = \frac{\det C}{\det C_{QF}}$$

or:

$$\operatorname{var}\mathscr{S}(d) = \frac{r^2\sigma_\varepsilon^2\sigma_r^2 + r^2\sin^2\varepsilon\sigma_\varepsilon^2\sigma_h^2 + \cos^2\sigma_r^2\sigma_h^2}{\sin^2\varepsilon\sigma_r^2 + r^2\cos^2\varepsilon\sigma_r^2 + \sigma_h^2} \tag{10.2}$$

which completes the calculation of the adjustment of the horizontal distance d.

Check. With $\Phi = h - r\sin\varepsilon = 0$ and $\Phi_X = |-\sin\varepsilon,\ -r\cos\varepsilon,\ 1|$ and

$$\mathscr{S}_X(d) = |(r^2\sigma_\varepsilon^2 + \sigma_h^2)\cos\varepsilon,\ -r\sin\varepsilon(\sigma_h^2 + \sigma_r^2),\ (\sigma_r^2 - r^2\sigma_\varepsilon^2)\cos\varepsilon\sin\varepsilon|$$

the checking formula is

$$\begin{aligned}\mathscr{S}_X(d)S\Phi_X^T = &-(r^2\sigma_\varepsilon^2\sigma_r^2+\sigma_h^2\sigma_r^2)\sin\varepsilon\cos\varepsilon+\\&+(\sigma_h^2\sigma_\varepsilon^2+\sigma_r^2\sigma_\varepsilon^2)r^2\cos\varepsilon\sin\varepsilon+\\&+(\sigma_r^2\sigma_h^2-r^2\sigma_\varepsilon^2\sigma_h^2)\cos\varepsilon\sin\varepsilon=0\ (!)\,.\end{aligned}$$

In fig. 3 part of the horizontal projection of the balloon's trajectory is shown.

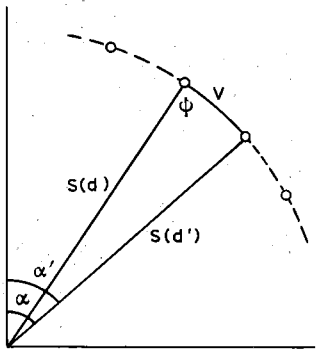

Fig. 3. Horizontal projection of the balloon's trajectory.

When in two successive positions (d, α) and (d', α') the corrected distances are taken, then the wind velocity and wind direction are derived from:

$$v = \frac{1}{\tau}\{\mathscr{S}^2(d)+\mathscr{S}^2(d')-2\,\mathscr{S}(d)\,\mathscr{S}(d')\cos(\alpha'-\alpha)\}^{\frac{1}{2}} \tag{10.3}$$

$$\psi = \arcsin\left\{\frac{\mathscr{S}(d')}{v\tau}\sin(\alpha'-\alpha)\right\} \tag{10.4}$$

where τ is the time interval.
A measure of precision for the wind is the standard vector error σ:

$$\sigma = \frac{\sqrt{2}}{\tau}\{\text{var}\ \mathscr{S}(d)+\mathscr{S}^2(d)\sigma_\alpha^2\}^{\frac{1}{2}}\,. \tag{10.5}$$

The wind being known, it is necessary to find the height for which the wind applies. In the system of radar radiosonde observations the height is measured twice, *viz.* by means of radar:

$$h_{\text{radar}} = r\sin\varepsilon$$

and independently by means of the radiosonde:

$$h_{\text{sonde}}$$

Therefore the measurement of the height pertains to the special class of direct "*unconditional*" observations, so that the "*best*" value for h is the weighted mean of both h_{radar} and h_{sonde} (*cf.* section 8).
One has

$$S = \begin{vmatrix} \sigma^2_{h_{\text{sonde}}} & 0 \\ 0 & \sigma^2_{h_{\text{radar}}} \end{vmatrix}$$

but

$$\sigma^2_{h_{\text{sonde}}} = \sigma^2_h$$

and according to (6.3)

$$\sigma^2_{h_{\text{radar}}} = \sin^2 \varepsilon \sigma^2_r + r^2 \cos^2 \varepsilon \sigma^2_\varepsilon$$

The weighted mean value becomes

$$\mathscr{S}(h) = \frac{\sigma^2_{h_{\text{sonde}}}}{\sigma^2_{h_{\text{sonde}}} + \sigma^2_{h_{\text{radar}}}} \boldsymbol{h}_{\text{radar}} + \frac{\sigma^2_{h_{\text{radar}}}}{\sigma^2_{h_{\text{sonde}}} + \sigma^2_{h_{\text{radar}}}} \boldsymbol{h}_{\text{sonde}}$$

or:

$$\mathscr{S}(h) = \frac{\sigma^2_h}{\sin^2 \varepsilon \sigma^2_r + r^2 \cos^2 \varepsilon \sigma^2_\varepsilon + \sigma^2_h} \boldsymbol{h}_{\text{radar}} + \frac{\sin^2 \varepsilon \sigma^2_r + r^2 \cos^2 \varepsilon \sigma^2_\varepsilon}{\sin^2 \varepsilon \sigma^2_r + r^2 \cos^2 \varepsilon \sigma^2_\varepsilon + \sigma^2_h} \boldsymbol{h}_{\text{sonde}}. \quad (10.6)$$

In addition the minimum variance is, see (8.18):

$$\text{var } \mathscr{S}(h) = \frac{1}{\frac{1}{\sigma^2_{h_{\text{sonde}}}} + \frac{1}{\sigma^2_{h_{\text{radar}}}}}$$

or:

$$\text{var } \mathscr{S}(h) = \frac{\sin^2 \varepsilon \sigma^2_r \sigma^2_h + r^2 \cos^2 \varepsilon \sigma^2_h \sigma^2_\varepsilon}{\sin^2 \varepsilon \sigma^2_r + r^2 \cos^2 \varepsilon \sigma^2_\varepsilon + \sigma^2_h}. \quad (10.7)$$

The problem has been solved and it can be supplied to an electronic computer. The result will depend on the errors σ^2_r, σ^2_ε and σ^2_h i.e. on the characteristics of the radar equipment, resp. the type of radiosonde used.
It should be remarked that the formulae involve a dosed mixture of true values and measured data, so that in practice the values of r, ε and h should be replaced by $\boldsymbol{r}$, $\boldsymbol{\varepsilon}$ and $\boldsymbol{h}$.

10.2. *Evaporation measurement*

Evaporation from a land surface can be measured by application of both the method of the energy balance and the aerodynamic method, *cf.* Priestley [1959].
In the aerodynamic method an expression for the evaporation is found which

reads

$$f_1 = C_1 \frac{\mathrm{d}\bar{u}}{\mathrm{d}\ln z}\left\{\left(\frac{\mathrm{d}e_s}{\mathrm{d}T}+\gamma\right)\frac{\mathrm{d}\bar{T}_w}{\mathrm{d}\ln z}-\gamma\frac{\mathrm{d}\bar{T}}{\mathrm{d}\ln z}\right\}. \tag{10.8}$$

According to the energy balance method f takes the form:

$$f_2 = C_2(H-S)\left\{1-\frac{\dfrac{\mathrm{d}\bar{T}}{\mathrm{d}\ln z}}{\left(\dfrac{\mathrm{d}e_s}{\mathrm{d}T}+\gamma\right)\dfrac{\mathrm{d}\bar{T}_w}{\mathrm{d}\ln z}}\right\}. \tag{10.9}$$

Here

T denotes the dry bulb temperature,
T_w the wet bulb temperature,
γ the psychrometric constant
$\frac{\mathrm{d}\bar{u}}{\mathrm{d}\ln z}$ the logarithmic windshear of the time average of the (horizontal) wind component,
$\frac{\mathrm{d}e_s}{\mathrm{d}T}$ the slope of the saturation vapour pressure curve,
$\frac{\mathrm{d}\bar{T}}{\mathrm{d}\ln z}$, respectively $\frac{\mathrm{d}\bar{T}_w}{\mathrm{d}\ln z}$ the logarithmic gradient of the time average of dry bulb, resp. wet bulb temperature,
H the net radiation of the surface, and
S the heat flux in the soil.

Both expressions (10.8) and (10.9) are equivalent functions in the sense of the present adjustment theory; they have some variables in common.
For convenience put

$$x_1 = C_1 \frac{\mathrm{d}\bar{u}}{\mathrm{d}\ln z}$$

$$x_2 = \gamma \frac{\mathrm{d}\bar{T}}{\mathrm{d}\ln z}$$

$$x_3 = \left(\frac{\mathrm{d}e_s}{\mathrm{d}T}+\gamma\right)\frac{\mathrm{d}\bar{T}_w}{\mathrm{d}\ln z}$$

$$x_4 = C_2(H-S)$$

Then both equations (10.8) and (10.9) can be written:

$$f_1 = x_1(x_3-x_2) \tag{10.10}$$

$$f_2 = \frac{x_4}{x_3}(x_3-x_2) \tag{10.11}$$

Here the solution of the problem of adjustment is restricted to the case that the base variables $x_1, x_2, \ldots x_4$ are again (stochastically) independent, so that the covariance matrix S is a diagonal matrix:

$$S = \begin{vmatrix} \sigma_1^2 & 0 \ldots 0 \\ 0 & \sigma_2^2 \quad \vdots \\ \vdots & \ddots \\ 0 \ldots & \sigma_4^2 \end{vmatrix}$$

Obviously the conditional equation takes a simple form. It can be derived from equating (10.10) and (10.11), which results into

$$\Phi = x_4 - x_1 x_3 = 0 \qquad (x_2 \neq x_3), \tag{10.12}$$

which after substituting for the original meaning of x_1, x_3 and x_4 takes the form:

$$C_2(H-S) = C_1 \frac{\mathrm{d}\bar{u}}{\mathrm{d}\ln z}\left(\frac{\mathrm{d}e_s}{\mathrm{d}T} + \gamma\right)\frac{\mathrm{d}\bar{T}_w}{\mathrm{d}\ln z}. \tag{10.13}$$

This formula is merely a statement of the law of conservation of energy for this problem. So this is an example of a conditional equation originating from a physical law. Due to the simple form of Φ, the Φ-approach is chosen. Referring to the scheme in section 9 one obtains

Variables:

$$x_1, \ldots x_4$$

Condition:

$$\Phi = x_4 - x_1 x_3 = 0$$

Function to be evaluated:

$$f = x_1(x_3 - x_2)$$

Jacobian:

$$\Phi_X = |-x_3, 0, -x_1, 1|$$

Covariance matrix:

$$\Gamma = \Phi_X S \Phi_X^T = x_3^2\sigma_1^2 + x_1^2\sigma_3^2 + \sigma_4^2$$

Gradient:

$$f_X = |x_3 - x_2, -x_1, x_1, 0|$$

Mixed covariance matrix:

$$f_X S\Phi_X^T = -x_3(x_3-x_2)\sigma_1^2 - x_1^2\sigma_3^2$$

$$f_X S\Phi_X^T \Gamma^{-1} = \frac{-x_3(x_3-x_2)\sigma_1^2 - x_1^2\sigma_3^2}{x_3^2\sigma_1^2 + x_1^2\sigma_3^2 + \sigma_4^2}$$

$$f_X S\Phi_X^T \Gamma^{-1}\Phi(X) = \frac{x_3(x_3-x_2)\sigma_1^2 + x_1^2\sigma_3^2}{x_3^2\sigma_1^2 + x_1^2\sigma_3^2 + \sigma_4^2}(x_1x_3 - x_4)$$

The estimator $\mathscr{S}(f)$ then takes the form

$$\mathscr{S}(f) = x_1(x_3-x_2) - \frac{x_3(x_3-x_2)\sigma_1^2 + x_1^2\sigma_3^2}{x_3^2\sigma_1^2 + x_1^2\sigma_3^2 + \sigma_4^2}(x_1x_3 - x_4). \tag{10.14}$$

It is observed that in the final outcome for the estimation the variance σ_2^2, i.e. the error of the dry bulb temperature gradient plays no part. The true values of x_i being unknown, one can find an approximation for $\mathscr{S}(f)$ when substituting the measured data for x_i. Then the expression can be rewritten in a more intelligible form:

$$\mathscr{S}(f) = \frac{x_3x_4(x_3-x_2)\sigma_1^2 + x_1^2(x_4 - x_1x_2)\sigma_3^2 + x_1(x_3-x_2)\sigma_4^2}{x_3^2\sigma_1^2 + x_1^2\sigma_3^2 + \sigma_4^2} \tag{10.15}$$

To calculate var $\mathscr{S}(f)$ write (10.14):

$$\mathscr{S}(f) = x_1(x_3-x_2) + \Lambda(x_4 - x_1x_3), \quad \Lambda = \frac{x_3(x_3-x_2)\sigma_1^2 + x_1^2\sigma_3^2}{x_3^2\sigma_1^2 + x_1^2\sigma_3^2 + \sigma_4^2}.$$

Then

$$\mathscr{S}_X(f) = |x_3 - x_2 - \Lambda x_3,\ -x_1,\ x_1 - \Lambda x_1,\ \Lambda|$$

and consulting the scheme:

$$\text{var } \mathscr{S}(f) = \mathscr{S}_X(f) S f_X^T,$$

$$\text{var } \mathscr{S}(f) = |x_3(1-\Lambda) - x_2,\ -x_1,\ x_1(1-\Lambda),\ \Lambda| \begin{vmatrix} \sigma_1^2 & 0 \ldots 0 \\ 0 \cdot & \\ \vdots & \ddots \\ 0 \cdots & \sigma_4^2 \end{vmatrix} \begin{vmatrix} x_3 - x_2 \\ -x_1 \\ x_1 \\ 0 \end{vmatrix}$$

which after some elementary computations reads:

$$\text{var } \mathscr{S}(f) = \frac{x_4^2\sigma_1^2\sigma_2^2 + x_1^2x_2^2\sigma_1^2\sigma_3^2 + (x_3-x_2)^2\sigma_1^2\sigma_4^2 + x_1^4\sigma_2^2\sigma_3^2 + x_1^2\sigma_2^2\sigma_4^2 + x_1^2\sigma_3^2\sigma_4^2}{x_3^2\sigma_1^2 + x_1^2\sigma_3^2 + \sigma_4^2} \tag{10.16}$$

Here the variance σ_2^2 is not lacking as occurred in $\mathscr{S}(f)$. After putting the real meaning of x_i in $\mathscr{S}(f)$ and var $\mathscr{S}(f)$ the problem of adjustment for measurements of evaporation is solved.

10.3. *Double theodolite observations*

In upper air research measurements are ofted made by using two (or more) theodolites, photo-theodolites or radiogoniometers that are tracking an airborne target from different sites. By simultaneously taking readings of the azimuthal angles and elevations the position of the target can be computed enabling the evaluation of the structure of upper air currents with high accuracy. Moreover, double theodolite observations are a tool for testing new radiosonde equipment, tracking constant level balloons and, in case of the use of photo-theodolites, for making stereoscopic pictures of noctilucent clouds and other cloud systems. Whereas the position of the target is determined by three coordinates, the two theodolites give four. The system is therefore overdetermined, double theodolite observations thus being an example of conditional observations. In fact, the four angles satisfy a certain relation.

Here special attention is paid to the adjustment of the height of the target. Making an analysis of the system, fig. 4 shows the details for determination

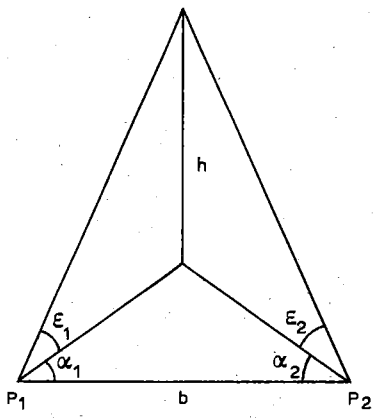

Fig. 4. Scheme for double theodolite observation.

of the height-formulae. The theodolites are situated in the points P_1 and P_2. The ground elevations of both points are assumed to be equal. Let b be the basis length $\overline{P_1 P_2}$. Some simple goniometric relations then give for the height:

$$h_1 = b \frac{\sin \alpha_1}{\sin(\alpha_1 + \alpha_2)} \tan \varepsilon_2 \,. \tag{10.17}$$

Another height formula is obviously

$$h_2 = b \frac{\sin\alpha_2}{\sin(\alpha_1+\alpha_2)} \tan\varepsilon_1 . \tag{10.18}$$

Both h_1 and h_2 involve three of the four coordinates $\alpha_1, \varepsilon_1, \alpha_2$ and ε_2. It is evident that two other height formulae must exist which contain the set of coordinates α_2, ε_1 and ε_2, resp. α_1, ε_1 and ε_2. It is easily to be derived that these take the form:

$$h_3 = b \tan\varepsilon_1 \tan\varepsilon_2 \left\{ \frac{\cos\alpha_2 \tan\varepsilon_1 \pm (\tan^2\varepsilon_2 - \tan^2\varepsilon_1 + \tan^2\varepsilon_1 \cos^2\alpha_2)^{\frac{1}{2}}}{\tan^2\varepsilon_1 - \tan^2\varepsilon_2} \right\}$$

and

$$h_4 = b \tan\varepsilon_1 \tan\varepsilon_2 \left\{ \frac{\cos\alpha_1 \tan\varepsilon_2 \pm (\tan^2\varepsilon_1 - \tan^2\varepsilon_2 + \tan^2\varepsilon_2 \cos^2\alpha_1)^{\frac{1}{2}}}{\tan^2\varepsilon_2 - \tan^2\varepsilon_1} \right\}$$

Although these expressions give identical values for exact values of the coordinates, their values may differ considerably when measured data are substituted, especially when the target is in the vicinity of the base plane through the points of observation P_1 and P_2. Then the accuracy of h_1 and h_2 is much inferior to the accuracy of h_3 and h_4. In the base plane h_1 and h_2 even become indeterminate. This feature is characteristic for a great number of problems. The phenomenon is to be attributed to the influence of the form of h in certain sub-areas of the working domain. The working area can be subdivided into regions where that height formula is to be taken that shows the smallest variance, *i.e.* for which the variance is min(var h_1, var h_2, var h_3, var h_4). So, when for any reason one can only use three of the four angles, then one can consult such a subdivision. When, however, all four coordinates are measured and recorded then the difficulties are overcome by application of the adjustment procedure. In fact, the four height formulae are merely four functions out of the infinite set of equivalent functions of the system of double theodolite observations. It is noted that all formulae in this class of equivalent height formulae, except for h_1, h_2, h_3 and h_4, are depending on all four coordinates. It has been supposed that the value of b is known with neglectable inaccuracy. The conditional equation is easily to be derived here by equating (10.17) and (10.18):

$$\Phi = \sin\alpha_1 \tan\varepsilon_2 - \sin\alpha_2 \tan\varepsilon_1 = 0 .$$

For reasons of symmetry the f-approach will be chosen. As the computations sometimes turn out to be laborious only the results are presented. In the scheme of the f-approach one has

Variables:

$$x_1 = \alpha_1 , \quad x_2 = \alpha_2 , \quad x_3 = \varepsilon_1 , \quad x_4 = \varepsilon_2$$

Covariance matrix:

$$S = \begin{vmatrix} \sigma_1^2 & 0 & \dots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \dots & & \sigma_4^2 \end{vmatrix}$$

There is only one condition, the number of equivalent height formulae to be taken therefore is *two*.
Set of equivalent height formulae:

$$F = \begin{vmatrix} b \dfrac{\sin\alpha_1}{\sin(\alpha_1+\alpha_2)} \tan\varepsilon_2 \\ b \dfrac{\sin\alpha_2}{\sin(\alpha_1+\alpha_2)} \tan\varepsilon_1 \end{vmatrix}$$

Jacobian

$$F_X = \frac{b}{\sin^2(\alpha_1+\alpha_2)} \times$$

$$\times \begin{vmatrix} \sin\alpha_2 \tan\varepsilon_2 & -\sin\alpha_1 \cos(\alpha_1+\alpha_2)\tan\varepsilon_2 & 0 & \sin\alpha_1 \sin(\alpha_1+\alpha_2)\sec^2\varepsilon_2 \\ -\sin\alpha_2 \cos(\alpha_1+\alpha_2)\tan\varepsilon_1 & \sin\alpha_1 \tan\varepsilon_1 & \sin\alpha_2 \sin(\alpha_1+\alpha_2)\sec^2\varepsilon_1 & 0 \end{vmatrix}$$

Further

$$QF_X = \frac{b}{\sin(\alpha_1+\alpha_2)} \left| \cos\alpha_1 \tan\varepsilon_2 \;\; -\cos\alpha_2 \tan\varepsilon_1 \;\; -\sin\alpha_2 \sec^2\varepsilon_1 \;\; \sin\alpha_1 \sec^2\varepsilon_2 \right| .$$

and

$$QC = QF_X S F_X^T = \frac{b^2}{\sin^3(\alpha_1+\alpha_2)} \times$$

$$\times \left| \begin{array}{l} \sin\alpha_2 \cos\alpha_1 \tan^2\varepsilon_2 \sigma_1^2 + \sin\alpha_1 \cos\alpha_2 \cos(\alpha_1+\alpha_2)\tan\varepsilon_1 \tan\varepsilon_2 \sigma_2^2 + \sin^2\alpha_1 \sin(\alpha_1+\alpha_2)\sec^4\varepsilon_2 \sigma_4^2 , \\ -\sin\alpha_2 \cos\alpha_1 \cos(\alpha_1+\alpha_2)\tan\varepsilon_1 \tan\varepsilon_2 \sigma_1^2 - \sin\alpha_1 \cos\alpha_2 \tan^2\varepsilon_1 \sigma_1^2 - \sin^2\alpha_2 \sin(\alpha_1+\alpha_2)\sec^4\varepsilon_1 \sigma_3^2 \end{array} \right| .$$

$$C_{QF} = QF_X S F_X^T Q =$$

$$= \frac{b^2}{\sin^2(\alpha_1+\alpha_2)} \left(\cos^2\alpha_1 \tan^2\varepsilon_2 \sigma_1^2 + \cos^2\alpha_2 \tan^2\varepsilon_1 \sigma_2^2 + \sin^2\alpha_2 \sec^4\varepsilon_1 \sigma_3^2 + \sin^2\alpha_1 \sec^4\varepsilon_2 \sigma_4^2\right) .$$

Putting the first element of $QC = t_1$, the second t_2 one obtains for C_F

$$C_F = \begin{vmatrix} F(x) \\ QF \end{vmatrix} = \begin{vmatrix} \boldsymbol{h}_1 & \boldsymbol{h}_2 \\ t_1 & t_2 \end{vmatrix}$$

Finally leading to the estimation

$$\mathscr{S}(h) = -\frac{\det C_F}{\det C_{QF}}$$

which has the complicated form:

$$\mathscr{S}(h) =$$
$$= \frac{\sin\alpha_2\cos\alpha_1\cos(\alpha_1+\alpha_2)\tan\varepsilon_1\tan\varepsilon_2\sigma_1^2+\sin\alpha_1\cos\alpha_2\tan^2\varepsilon_1\sigma_2^2+\sin^2\alpha_2\sin(\alpha_1+\alpha_2)\sec^4\varepsilon_1\sigma_3^2}{\sin(\alpha_1+\alpha_2)(\cos^2\alpha_1\tan^2\varepsilon_2\sigma_1^2+\cos^2\alpha_2\tan^2\varepsilon_1\sigma_2^2+\sin^2\alpha_2\sec^4\varepsilon_1\sigma_3^2+\sin^2\alpha_1\sec^4\varepsilon_2\sigma_4^2}\boldsymbol{h}_1$$
$$+\frac{\sin\alpha_2\cos\alpha_1\tan^2\varepsilon_1\sigma^2+\sin\alpha_1\cos\alpha_2\cos(\alpha_1+\alpha_2)\tan\varepsilon_1\tan\varepsilon_2\sigma_2^2+\sin^2\alpha_1\sin(\alpha_1+\alpha_2)\sec^4\varepsilon_2\sigma_4^2}{\sin(\alpha_1+\alpha_2)(\cos^2\alpha_1\tan^2\varepsilon_2\sigma_1^2+\cos^2\alpha_2\tan^2\varepsilon_1\sigma_2^2+\sin^2\alpha_2\sec^4\varepsilon_1\sigma_3^2+\sin^2\alpha_1\sec^4\varepsilon_2\sigma_4^2}\boldsymbol{h}_2 \tag{10.19}$$

The expression again involves terms with true values and measured data. In practice the coefficients of $\boldsymbol{h}_1$ and $\boldsymbol{h}_2$ are to be approximated by replacing the true values by measured data. When such a substitution is carried out, then the form of $\mathscr{S}(h)$ can be changed by rearranging terms to obtain

$$\mathscr{S}(h) = \frac{(\cos^2\alpha_2\tan\varepsilon_1\sigma_2^2+\sin^2\alpha_2\sec^4\varepsilon_1\sigma_3^2)\boldsymbol{h}_1+(\cos^2\alpha_1\tan^2\varepsilon_2\sigma_1^2+\sin^2\alpha_1\sec^4\varepsilon_2\sigma_4^2)\boldsymbol{h}_2}{\cos^2\alpha_1\tan^2\varepsilon_2\sigma^2+\cos^2\alpha_2\tan^2\varepsilon_1\sigma_2^2+\sin^2\alpha_2\sec^4\varepsilon_1\sigma_3^2+\sin^2\alpha_1\sec^4\varepsilon_2\sigma_4^2} \tag{10.20}$$

The minimum variance is determined from the formula $\operatorname{var}\mathscr{S}(h) = \dfrac{\det C}{\det C_{QF}}$.
The denumerator $\det C_{QF}$ is already known, the numerator is to be evaluated by means of the (2, 2) minors of F_X.
The terms are successively:

$$\sigma_1^2\sigma_2^2: \quad \frac{b^4}{\sin^8(\alpha_1+\alpha_2)}\tan^2\varepsilon_1\tan^2\varepsilon_2\sin^2\alpha_1\sin^2\alpha_2\{1-\cos^2(\alpha_1+\alpha_2)\} =$$
$$= \frac{b^4}{\sin^6(\alpha_1+\alpha_2)}\tan^2\varepsilon_1\tan^2\varepsilon_2\sin^2\alpha_1\sin^2\alpha_2$$

$$\sigma_1^2\sigma_3^2: \quad \frac{b^4}{\sin^8(\alpha_1+\alpha_2)}\sin^4\alpha_2\sin^2(\alpha_1+\alpha_2)\tan^2\varepsilon_2\sec^4\varepsilon_1 =$$
$$= \frac{b^4}{\sin^6(\alpha_1+\alpha_2)}\sin^4\alpha_2\tan^2\varepsilon_2\sec^4\varepsilon_1$$

$$\sigma_1^2\sigma_4^2: \quad \frac{b^4}{\sin^8(\alpha_1+\alpha_2)}\sin^2\alpha_1\sin^2\alpha_2\sin^2(\alpha_1+\alpha_2)\cos^2(\alpha_1+\alpha_2)\tan^2\varepsilon_1\sec^4\varepsilon_2 =$$
$$= \frac{b^4}{\sin^6(\alpha_1+\alpha_2)}\sin^2\alpha_1\sin^2\alpha_2\cos^2(\alpha_1+\alpha_2)\tan^2\varepsilon_1\sec^4\varepsilon_2$$

$$\sigma_2^2\sigma_3^2: \quad \frac{b^4}{\sin^8(\alpha_1+\alpha_2)}\sin^2\alpha_1\sin^2\alpha_2\sin^2(\alpha_1+\alpha_2)\cos(\alpha_1+\alpha_2)\tan^2\varepsilon_2\sec^4\varepsilon_1 =$$

$$= \frac{b^4}{\sin^6(\alpha_1+\alpha_2)}\sin^2\alpha_1\sin^2\alpha_2\cos(\alpha_1+\alpha_2)\tan^2\varepsilon_2\sec^4\varepsilon_1$$

$$\sigma_2^2\sigma_4^2: \quad \frac{b^4}{\sin^8(\alpha_1+\alpha_2)}\sin^2\alpha_1\sin^2(\alpha_1+\alpha_2)\tan^2\varepsilon_1\sec^4\varepsilon_2 =$$

$$= \frac{b^4}{\sin^6(\alpha_1+\alpha_2)}\sin^2\alpha_1\tan^2\varepsilon_1\sec^4\varepsilon_2$$

$$\sigma_3^2\sigma_4^2: \quad \frac{b^4}{\sin^8(\alpha_1+\alpha_2)}\sin^2\alpha_1\sin^2\alpha_2\sin^4(\alpha_1+\alpha_2)\sec^4\varepsilon_1\sec^4\varepsilon_2 =$$

$$= \frac{b^4}{\sin^6(\alpha_1+\alpha_2)}\sin^2\alpha_1\sin^2\alpha_2\sin^2(\alpha_1+\alpha_2)\sec^4\varepsilon_1\sec^4\varepsilon_2$$

Hence, var $\mathscr{S}(h)$ becomes:

$$\text{var } \mathscr{S}(h) = \frac{b^2}{\sin^4(\alpha_1+\alpha_2)\{\cos^4\varepsilon_1(\sin^2\alpha_1\sigma_4^2+\cos^2\alpha_1\cos^2\varepsilon_2\sin^2\varepsilon_2\sigma_2^2)+\cos^4\varepsilon_2(\sin^2\alpha_2\sigma_3^2+\cos^2\alpha_2\sin^2\varepsilon_1\cos^2\varepsilon_1\sigma_1^2}$$
$$\times [\sin^2\alpha_1\sin^2\alpha_2\sin^2\varepsilon_1\sin^2\varepsilon_2\cos^2\varepsilon_1\cos^2\varepsilon_2\sin^2(\alpha_1+\alpha_2)\sigma_1^2\sigma_2^2+\sin^4\alpha_2\sin^2\varepsilon_2\cos^2\varepsilon_2\,\sigma_1^2\sigma_3^2$$
$$+ \sin^2\alpha_1\sin^2\alpha_2\cos^2(\alpha_1+\alpha_2)\cos^2\varepsilon_1\sin^2\varepsilon_1\,\sigma_1^2\sigma_4^2+\sin^2\alpha_1\sin^2\alpha_2\cos^2(\alpha_1+\alpha_2)\cos^2\varepsilon_2\sin^2\varepsilon_2\sigma_2^2\sigma_3^2$$
$$+ \sin^4\alpha_1\sin^2\varepsilon_1\cos^2\varepsilon_1\sigma_2^2\sigma_4^2+\sin^2\alpha_1\sin^2\alpha_2\sin^2(\alpha_1+\alpha_2)\sigma_3^2\sigma_4^2] . \qquad (10.21)$$

In order to have an impression of the variance reduction when applying the theory of adjustment, var $\mathscr{S}(h)$ has been evaluated for the special case that $\sigma_1=\sigma_2=\sigma_3=\sigma_4=0.001$ rad. These are reasonable values for the inaccuracy of presently used theodolites. Furthermore, the base length has been chosen to be 5 km. Figure 5 shows the isopleths of $\sigma=(\text{var } \mathscr{S}(h))^{\frac{1}{2}}$ (m) at a level 5 km above the ground surface. This graph should be compared with the graph for, say, the root mean square error of h_1. In fig. 6 one can find the result for $\sigma_{h_1}=\text{var}^{\frac{1}{2}}h_1$ under similar conditions as in fig. 5. It is to be remarked that σ_{h_1} is calculated by means of formula (6.3).

$$\sigma_{h_1}^2=\text{var } h_1=h_{1,X}Sh_{1,X}^T$$

$h_{1,X}$ is the first row of matrix F_X, so that

$$\sigma_{h_1} = \frac{b}{\sin^2(\alpha_1+\alpha_2)}\{\sin^2\alpha_2\tan^2\varepsilon_2\sigma_1^2+\sin^2\alpha_1\cos^2(\alpha_1+\alpha_2)\tan^2\varepsilon_2\sigma_2^2+\sin^2\alpha_1\sin^2(\alpha_1+\alpha_2)\sec^4\varepsilon_2\sigma_4^2\}^{\frac{1}{2}}$$

From both the figures 5 and 6 it is discernable that the reduction of variance with respect to the height formula h_1 is substantial in the vicinity of the base plane. So, when a balloon is launched near one of the points of observations

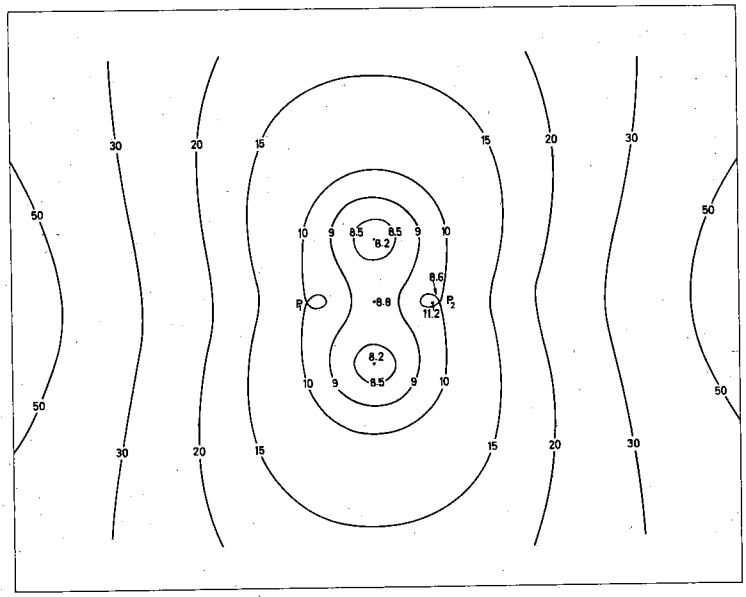

Fig. 5. Double theodolite observation. Isopleths of the (minimum) root mean square error σ (m) associated with the estimator $\mathscr{S}(h)$ at a level of 5 km above the ground surface. $\sigma_1=\sigma_2=\sigma_3=\sigma_4=0.001$ rad. Base length 5 km.

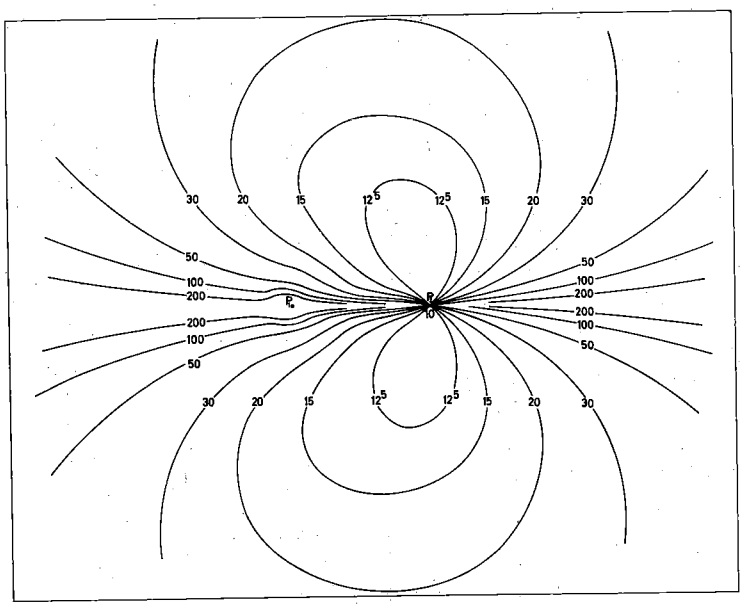

Fig. 6. Double theodolite observation. Isopleths of the root mean square error σ_{h_1} (m) associated with the conventional height formula h_1 at a level of 5 km above the ground surface. $\sigma_1=\sigma_2=\sigma_3=\sigma_4=0.001$ rad. Baselength $\overline{P_1P_2}=5$ km.

or near the baseline one should be very careful in choosing a height formula. This is especially true when the formula for adjustment is too complicated for mere desk calculation and an electronic computer is not available. To demonstrate the effect, the table below shows the height data of a balloon ascent during the first few minutes. A pilot balloon was launched near point P_1 and the heights were computed using h_1, h_2, h_3, h_4 and $\mathscr{S}(h)$. As could be expected the values of h_1 exceed those of the other formulae considerably. Those of h_2 are still reliable as the balloon was moving near point P_1 where the graph of σ_{h_2} will show a singular point. In fact, the graph of σ_{h_2} is the image of that of σ_{h_1} with respect to the line bisecting the baseline P_1P_2. Figure 7 gives a

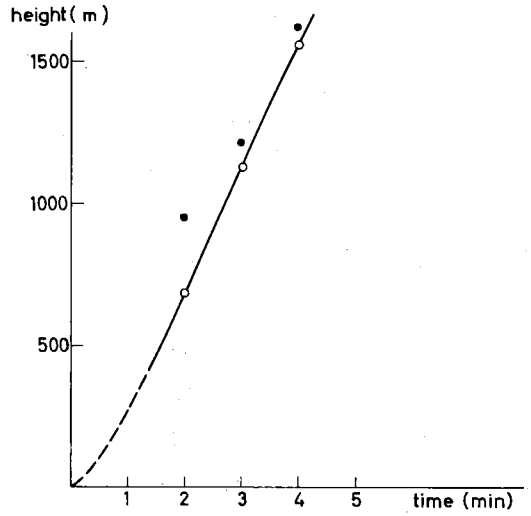

Fig. 7. Ascent curve for a pilot balloon determined from a double theodolite observation, a few minutes after release near one of the points of observation. Base length 4481 m.

graph of the ascent curve for the first couple of minutes. All height values of h_2, h_3 and h_4 are within the circles representing the data for $\mathscr{S}(h)$. Those of h_1 suggest a higher ascent rate in the first two minutes after launching and

TABLE

Balloon height versus time for a double theodolite observation

time (min.)	$h_1(\alpha_1, \alpha_2, \varepsilon_2)$	$h_2(\alpha_1, \alpha_2, \varepsilon_1)$	$h_3(\varepsilon_1, \varepsilon_2, \alpha_1)$	$h_4(\varepsilon_1, \varepsilon_2, \alpha_2)$	$\mathscr{S}(h)(\alpha_1, \alpha_2, \varepsilon_1, \varepsilon_2)$
2	953	679	690	690	681
3	1218	1127	1132	1146	1129
4	1623	1559	1561	1569	1562

thereafter a deceleration, which from a physical point of view is very doubtful. On the contrary, the ascent curve, which is based on the $\mathscr{S}(h)$ values, gives a more reliable impression of the ascent.
Examining figure 5 it is observed that minima are found at the line bisecting the base P_1P_2 at a considerable distance from the base. The minimum is not on the baseline proper as is sometimes thought. It can be shown that the minimum moves outward when the altitude increases. Apparently for a given base and height of an object there is an optimal distance to measure that height with highest precision. Furthermore it can be noted that at the base plane the minimum variance σ^2 equals the variance of both h_3 and h_4.

10.4. *Observation of rainfall with inclined rain gauges*
In a paper due to Levert [1962] the theory and practice are described of the observation of rainfall with gauges that are oblique with respect to the vertical. The gauges are installed in regular arrays and for several arrangements formulae are presented for the inclination, intensity and azimuthal angle of the rain as a function of collected amounts of precipitation and some parameters of the installation. In practically all arrangements conditional equations exist between the observed amounts of rain so that the theory of adjustment comes to the fore. Here the case is studied of four gauges being installed at the end of a rectangle cross, all equally inclined to the vertical and a fifth situated in the centre with no inclination at all (fig. 8).

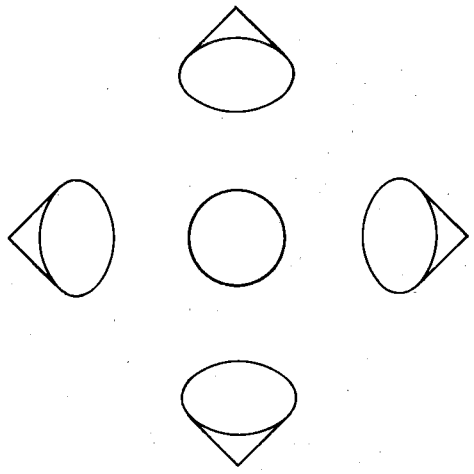

Fig. 8. Installation of five rain gauges to measure inclination, intensity and azimuth angle of rain.

When x_1, x_2, x_3 and x_4 represent the amounts of rain collected in the four outer gauges and x_0 the amount in the centre gauge, it can be demonstrated that the total amount of precipitation is:

$$f=x_0\,,$$

the azimuthal angle of the rain is given by:

$$\tan\psi = \frac{x_3-x_4}{x_1-x_2}$$

and the inclination of the rain to the vertical:

$$\tan\gamma = \frac{\{(x_1-x_2)^2+(x_3-x_4)^2\}^{\frac{1}{2}}}{2x_0\sin\gamma_g}$$

where γ_g represents the inclination of the gauges with respect to the vertical. It should be noted that the rainfall is considered to be uniformly distributed. The mutual distances of the gauges should therefore not be too excessive.
Between the amounts of rain $x_0, x_1, \ldots x_4$ the following conditions exist:

$$\varphi_1=x_1+x_2-x_3-x_4=0$$
$$\varphi_2=x_1+x_2-2x_0\cos\gamma_g=0\,.$$

Apparently the conditional equations are linear forms in the variables x_0, $x_1, \ldots x_4$. It therefore seems that the Φ-approach is preferable to the f-approach. It is supposed that the quantities $x_0, x_1, \ldots x_4$ are measured with equal accuracy, so that the covariance matrix S reads:

$$S=\sigma^2 E\,.$$

This is an example of more than one condition and also of more than one quantity to be adjusted. One could then adjust the base variables first and substitute the corrected values of $x_0, \ldots x_4$ in the quantities, but it is preferred here to adjust all three quantities separately.
Referring to the scheme one has

Variables:

$$x_0 \quad x_1 \quad x_2 \quad x_3 \quad x_4$$

Covariance matrix:

$$S=\sigma^2 E$$

Set of conditions:

$$\Phi = \begin{vmatrix} x_1+x_2-x_3-x_4 \\ x_1+x_2-ax_0 \end{vmatrix}$$

where

$$a=2\cos\gamma_g$$

Jacobian:

$$\Phi_X=\begin{vmatrix} 0 & 1 & 1 & -1 & -1 \\ -a & 1 & 1 & 0 & 0 \end{vmatrix}$$

Covariance matrix:

$$\Gamma=\Phi_X S\Phi_X^T=\sigma^2\Phi_X\Phi_X^T=\sigma^2\begin{vmatrix} 4 & 2 \\ 2 & a^2+2 \end{vmatrix}$$

The inverse of Γ:

$$\Gamma^{-1}=\frac{1}{4\sigma^2(a^2+1)}\begin{vmatrix} a^2+2 & -2 \\ -2 & 4 \end{vmatrix}$$

Next the amount of precipitation is calculated with the adjustment procedure. One has

$$f=x_0\,.$$

Gradient

$$f_X=\begin{vmatrix} 1 & 0 & 0 & 0 & 0 \end{vmatrix}$$

Mixed covariance matrix

$$f_X S\Phi_X^T=\sigma^2 f_X\Phi_X^T=\sigma^2\begin{vmatrix} 0 & -a \end{vmatrix}$$

Further

$$f_X S\Phi_X^T\Gamma^{-1}=\frac{1}{4(a^2+1)}\begin{vmatrix} 0 & -a \end{vmatrix}\begin{vmatrix} a^2+2 & -2 \\ -2 & 4 \end{vmatrix}=\frac{a}{2(a^2+1)}\begin{vmatrix} 1 & -2 \end{vmatrix}$$

and

$$f_X S\Phi_X^T\Gamma^{-1}\Phi(\boldsymbol{x})=\frac{a}{2(a^2+1)}\begin{vmatrix} 1 & -2 \end{vmatrix}\begin{vmatrix} \boldsymbol{x}_1+\boldsymbol{x}_2-\boldsymbol{x}_3-\boldsymbol{x}_4 \\ \boldsymbol{x}_1+\boldsymbol{x}_2-a\boldsymbol{x}_0 \end{vmatrix}=$$

$$=\frac{a}{2(a^2+1)}\left(2a\boldsymbol{x}_0-\sum_{i=1}^{4}\boldsymbol{x}_i\right)$$

to obtain

$$\mathscr{S}(f)=f(\boldsymbol{x})-f_X S\Phi_X\Gamma^{-1}\Phi(\boldsymbol{x})=\boldsymbol{x}_0-\frac{a}{2(a^2+1)}\left(2a\boldsymbol{x}_0-\sum_{i=1}^{4}\boldsymbol{x}_i\right)$$

and after substituting for a

$$\mathscr{S}(f) = \frac{x_0 + \cos\gamma_g \sum_{i=1}^{4} x_i}{4\cos^2\gamma_g + 1}.$$

To calculate var $\mathscr{S}(f)$ determine $\mathscr{S}_X(f)$:

$$\mathscr{S}_X(f) = \frac{1}{2(a^2+1)} \begin{vmatrix} 2 & a & a & a & a \end{vmatrix}$$

Then

$$\text{var } \mathscr{S}(f) = \mathscr{S}_X(f) S f_X^T = \sigma^2 \mathscr{S}_X(f) f_X^T = \frac{\sigma^2}{2(a^2+1)} \begin{vmatrix} 2 & a & a & a & a \end{vmatrix} \begin{vmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{vmatrix}$$

or

$$\text{var } \mathscr{S}(f) = \frac{\sigma^2}{4\cos^2\gamma_g + 1}.$$

For checking $\mathscr{S}(f)$ one should verify:

$$\mathscr{S}_X(f) S \Phi_X^T = 0$$

or

$$\sigma^2 \mathscr{S}_X(f) \Phi_X^T = \frac{\sigma^2}{2(a^2+1)} \begin{vmatrix} 2 & a & a & a & a \end{vmatrix} \begin{vmatrix} 0 & -a \\ 1 & 1 \\ 1 & 1 \\ -1 & 0 \\ -1 & 0 \end{vmatrix} = \begin{vmatrix} 0 & 0 \end{vmatrix}$$

which emphasizes the correctness of the solution to the problem.
Next one proceeds with the adjustment procedure for the azimuthal angle ψ.

$$f = \tan\psi = \frac{x_3 - x_4}{x_1 - x_2}$$

Gradient

$$f_X = \left| 0 \quad \frac{x_3 - x_4}{(x_1 - x_2)^2} - \frac{x_3 - x_4}{(x_1 - x_2)^2} \quad \frac{1}{x_1 - x_2} - \frac{1}{x_1 - x_2} \right|.$$

Mixed covariance matrix

$$f_X S \Phi_X^T = \sigma^2 f_X \Phi_X^T = 0,$$

showing that

$$\mathscr{S}(f)=f.$$

The original formula for f proves—more or less by accident—also to be the "*best*" in the sense that has been outlined in this study. The variance of f is at the same time the minimum variance:

$$\text{var } \mathscr{S}(f)=f_X S f_X^T=2\sigma^2\frac{(x_3-x_4)^2+(x_1-x_2)^2}{(x_1-x_2)^4}.$$

Finally, applying the theory of adjustment to the inclination of the rain consider

$$f=\tan\gamma=\frac{\{(x_1-x_2)^2+(x_3-x_4)^2\}^{\frac{1}{2}}}{x_0}$$

where $b=2\sin\gamma_g$.

Gradient:

$$f_X=\frac{1}{2f}\left|-\frac{2\{(x_1-x_2)^2+(x_3-x_4)^2\}}{bx_0^3}\;\frac{2(x_1-x_2)}{b^2x_0^2}\right.$$
$$\left.-\frac{2(x_1-x_2)}{b^2x_0^2}\;\frac{2(x_3-x_4)}{b^2x_0^2}-\frac{2(x_3-x_4)}{b^2x_0^2}\right|$$
$$=\frac{1}{bfx_0^2}\left|-b^2x_0f^2\;x_1-x_2\;\;x_2-x_1\;\;x_3-x_4\;\;x_4-x_3\right|$$

Mixed covariance matrix:

$$f_X S\Phi_X^T=\sigma^2 f_X\Phi_X^T=\sigma^2\begin{vmatrix}0 & \frac{af}{x_0}\end{vmatrix}$$

Further

$$f_X S\Phi_X^T\Gamma^{-1}=\frac{\sigma^2}{4\sigma^2(a^2+1)}\begin{vmatrix}0 & \frac{af}{x_0}\end{vmatrix}\begin{vmatrix}a^2+2 & -2\\ -2 & 4\end{vmatrix}=\frac{af}{2x_0(a^2+1)}\left|-1\quad 2\right|$$

and

$$f_X S\Phi_X^T\Gamma^{-1}\Phi(\boldsymbol{x})=\frac{af}{2x_0(a^2+1)}\left|-1\quad 2\right|\begin{vmatrix}\boldsymbol{x}_1+\boldsymbol{x}_2-\boldsymbol{x}_3-\boldsymbol{x}_4\\ \boldsymbol{x}_1+\boldsymbol{x}_2-a\boldsymbol{x}_0\end{vmatrix}=$$
$$=\frac{af}{2x_0(a^2+1)}\left(\sum_{i=1}^{4}\boldsymbol{x}_i-2a\boldsymbol{x}_0\right)$$

to obtain

$$\mathscr{S}(f)=f(\boldsymbol{x}_0,\ldots\boldsymbol{x}_4)+\frac{af}{2x_0(a^2+1)}\left(2a\boldsymbol{x}_0-\sum_{i=1}^{4}\boldsymbol{x}_i\right).$$

For practical use the true values are replaced by the measured data:

$$\mathscr{S}(f)=f(x_0,\dots x_4)\left[1+\frac{a}{2x_0(a^2+1)}\left(2ax_0-\sum_{i=1}^{4}x_i\right)\right]$$

and after substituting for a and b:

$$\mathscr{S}(\tan\gamma)=\frac{\{(x_1-x_2)^2+(x_3-x_4)^2\}^{\frac{1}{2}}}{2x_0\sin\gamma_g}\times$$

$$\times\left[1+\frac{\cos\gamma_g}{x_0(4\cos^2\gamma_g+1)}\left\{2x_0\cos\gamma_g-\sum_{i=1}^{4}x_i\right\}\right].$$

To calculate the minimum variance determine again $\mathscr{S}_X(f)$:

$$\mathscr{S}_X(f)=\frac{1}{b^2fx_0^2}\left|-b^2x_0f^2\quad x_1-x_2\quad x_2-x_1\quad x_3-x_4\quad x_4-x_3\right|$$

$$+\frac{af}{2x_0(a^2+1)}\left|2a\quad -1\quad -1\quad -1\quad -1\right|$$

Then

$$\text{var}\,\mathscr{S}(f)=\mathscr{S}_X(f)Sf_X^T=\sigma^2\mathscr{S}_X(f)\begin{vmatrix}-b^2x_0f^2\\ x_1-x_2\\ x_2-x_1\\ x_3-x_4\\ x_4-x_3\end{vmatrix}\frac{1}{b^2fx_0^2}$$

or

$$\text{var}\,\mathscr{S}(f)=\frac{\sigma^2}{b^4f^2x_0^4}\{f^4b^4x_0^2+2(x_1-x_2)^2+2(x_3-x_4)^2\}-\frac{\sigma^2a^2f^2}{x_0^2(a^2+1)}.$$

Finally after replacing a and b:

$$\text{var}\,\mathscr{S}(\tan\gamma)=\text{var}(\tan\gamma)-\frac{4\sigma^2\tan^2\gamma\cos^2\gamma_g}{x_0(4\cos^2\gamma_g+1)}.$$

It may be easily shown that with $\mathscr{S}_X(f)$ the checking formula

$$\mathscr{S}_X(f)S\Phi_X^T=0$$

is indeed satisfied. This completes the calculations for this type of measurement.

11. Use of high speed electronic computers

The calculation associated with the theory of adjustment is often laborious and cumbersome. This is particularly true when the number of conditions and consequently the order of the matrices increases. The matrices involve functions of one or more variables which make it clear that to obtain closed forms for the estimation and variance the rank of say Γ and C cannot be much higher than three or four. Moreover, the final results are often so complicated that their interpretation cannot in general be adapted to manual and graphical techniques and will be unsuitable in practice when quick results are required. However, much work can be saved when high speed electronic computers are employed. The merit of a computer program is not only its speed but also the possibility to handle a great number of data.

When for example the estimator (10.20) and variance (10.21) of double theodolite observations are examined it is relevant that mere hand calculation is so time-consuming that in the absence of an electronic computer application of the method in practice will be doubtfull. For this reason a program has been developed using the ZEBRA, a computer of moderate speed developed by the Postal and Telegraphic Service in the Netherlands. A balloon ascent involving 60 minute-observations takes less than 3 minutes computer-time. The time needed for printing the data amounts to approximately 30 minutes.

There is still another argument for emphasizing the extreme adequacy of the electronic computers for the wide class of problems here. Examining the scheme for matrix calculation of the estimator $\mathscr{S}(f)$ and var $\mathscr{S}(f)$ it is observed that the calculation starts after the Jacobian Φ_X resp. F_X for a special problem has been established. The remaining part is routine computation. It will therefore suffice to design a standard program for the process which proceeds with the calculation of the covariance matrices Γ resp. C. The first part of the process needs the development of a sub-program and includes the contents of S, Φ or F, Φ_X respectively F_X for the special problem under consideration.

When such a routine is followed it covers a very wide category of observations ranging from sample means to repeated conditional measurements. The routine also includes problems of regression analyses, parabolic interpolation and stochastically dependent observations.

It is a matter of machine capacity whether problems involving matrices of order 50 to 100 are still suitable to be handled by these high speed computers.

REFERENCES

[1948] Birkhoff, G., and Mac Lane, S., *A Survey of Modern Algebra.* New York.

[1959] Priestley, C. H. B., *Turbulent transfer in the lower atmosphere.* Chicago.

[1961] Linnik, Y. V., *Method of Least Squares and Principles of the Theory of Observations.* (Translated from the Russian). London-New York-Paris.
(French and German translations also available).

[1961] Ritsema, A. R., and Scholte, J. G. J., Note on the determination of the best-fitting plane for a given set of directions. *Geofis. pura e appl.* **49**, 13–14.

[1962] Levert, C., Praxis und Theorie der Messung mit schrägen Regenmessern. *Archiv für Meteorologie, Geophysik und Bioklimatologie* **11**. 447–467.

[1934] Kähler, E., *Einführung in die Theorie der Systeme von Differentialgleichungen.* Hamburger Mathem. Einzelschriften **16**. Leipzig.
(Reprinted together with Blaschke, W., Vorlesungen über Integralgeometrie, New York).

SUMMARY

A study is presented on the effect of replacing the set of (non-linear) conditional equations used in the method of least squares for the problem of adjustment of observations by an equivalent set and on related problems. It is shown that the replacement eventually does not change the results of the adjustment procedure.

The method of investigation leads in a natural way to a new approach to the adjustment problem in which no conditional equations occur but equivalent formulae for the quantities to be adjusted. The new method is particullary suitable to be programmed for high speed electronic computors and possesses a flexibility which substantially facilitates its application especially in the case of non-linear conditional equations.

The general theory is applied to four examples of adjustment problems in meteorology including the problem which gave the impetus for the present study.

SAMENVATTING

Onderzocht werden de gevolgen van de vervanging van een stelsel van (niet-lineaire) voorwaardenvergelijkingen in de methode der kleinste kwadraten voor het probleem van de vereffening van waarnemingen door een gelijkwaardig stelsel en verdere hiermeden samenhangende problemen. Aangetoond werd dat de vervanging geen enkele invloed heeft op het uiteindelijke resultaat van de vereffening.
De gevolgde methode van onderzoek leidde op natuurlijke wijze tot een nieuwe aanpak van het vereffeningsprobleem waarin geen voorwaardenvergelijkingen meer voorkomen maar een of meer gelijkwaardige formules voor de te vereffenen grootheden. De nieuwe methode is in het bijzonder geschikt om geprogrammeerd te worden voor electronische rekenmachines en bezit verder een groot aanpassingsvermogen hetgeen de toepassing vergemakkelijkt speciaal voor het geval dat de voorwaardenvergelijkingen niet lineair zijn.
De algemene theorie werd toegepast op een viertal voorbeelden uit de meteorologie waaronder het probleem dat de aanleiding tot onderzoek vormde.

Van de reeks MEDEDELINGEN EN VERHANDELINGEN *zijn bij het Staatsdrukkerij- en Uitgeverijbedrijf nog verkrijgbaar de volgende nummers:*

23, 25, 26, 27, 29*b*, 30, 31, 34*b*, 36, 37, 38, 39, 40, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59.

alsmede:

60. C. Kramer, J. J. Post en J. P. M. Woudenberg. Nauwkeurigheid en betrouwbaarheid van temperatuur- en vochtigheidsbepalingen in buitenlucht met behulp van kwikthermometers. 1954. (60 blz., 11 fig.) . . . . . f 3,50

62. C. Levert. Regens. Een statistische studie. 1954. (246 blz., 67 fig., 143 tab.) 10,—

63. P. Groen. On the behaviour of gravity waves in a turbulent medium, with application to the decay and apparent period increase of swell. 1954. (23 blz.) 1,50

64. H. M. de Jong. Theoretical aspects of aeronavigation and its application in aviation meteorology. 1956. (124 blz., 80 fig., 9 krt. en tab.) 4.50

65. J. G. J. Scholte. On seismic waves in a spherical earth. 1956. (55 blz., 24 fig.) 5,—

66. G. Verploegh. The equivalent velocities for the Beaufort estimates of the wind force at sea. 1956. (38 blz., 17 tab.) 1,75

67. G. Verploegh. Klimatologische gegevens van de Nederlandse lichtschepen over de periode van 1910–1940.

Dl. 1: Stormstatistieken.—Climatological data of the Netherlands light-vessels over the period 1910–1940. P. 1: Statistics of gales. 1956. (68 blz. en tab.) 3,50

Dl. 2: Luchtdruk en wind; zeegang.—Climatological data of the Netherlands light-vessels over the period 1910–1940. P. 2: Airpressure and wind; state of the sea. 1958. (91 blz. en tab.) 7,50

Dl. 3: Temperaturen en hydrometeoren; onweer.—Climatological data of the Netherlands light-vessels over the period 1910–1940. P. 3: Temperatures and hydrometeors; thunderstorms. 1959. (146 blz. en tab.) . . . . . 8,—

68. F. H. Schmidt. On the diffusion of stack gases in the atmosphere. 1957. (60 blz., 12 fig. en tab.) . . . . . 5,—

69. H. P. Berlage. Fluctuations of the general atmospheric circulation of more than one year; their nature and prognostic value. 1957. (152 blz., 35 fig. en 33 tab.) 7,50

70. C. Kramer. Berekening van de gemiddelde grootte van de verdamping voor verschillende delen van Nederland volgens de methode van Penman. 1957. (85 blz., 15 fig. en tab.) . . . . . 7,—

71. H. C. Bijvoet. A new overlay for the determination of the surface wind over sea from surface weather charts. 1957. (35 blz., 15 fig. en 8 tab.) 2,50

72. J. G. J. Scholte. Rayleigh waves in isotropic and anisotropic elastic media. 1958. (43 blz., 12 fig. en tab.) . . . . . 3,—

73. M. P. H. Weenink. A theory and method of calculation of wind effects on sea levels in a partly-enclosed sea, with special application to the southern coast of the North Sea. 1958. (111 blz., 28 fig. en 2 tab.) . . . . . 8,—

74. H. M. de Jong. Geostrophic flow. Geostrophic approximation in the upper air flow with application to aeronavigation and air trajectories. 1959. (100 blz. 17 fig., 14 krt. en 2 tab.) . . . . . 5,—

75. S. W. Visser. A new analysis of some features of the 11-year and 27-day cycles in solar activity and heir reflection in geophysical phenomena. 1959. (65 blz., 16 fig. en 12 tab.) 3,50

76. A. R. Ritsema and J. Veldkamp. Fault plane mechanisms of South East Asian earthquakes. 1960. (63 blz., 26 fig. en 11 tab.) . . . . . 4,—

77. G. Verploegh. On the annual variation of climatic elements of the Indian Ocean. P. 1: Text. P. 2: Charts. 1960. (64 blz., 15 fig. en 28 krt.) . . . . . 6,—

78. J. A. As. Instruments and measuring methods in paleomagnetic research. 1960. (56 blz., 20 fig.) 2,50

79. D. J. Bouman. Consistency of approximations in discontinuous fields of motion in the atmosphere with an introduction to the use of generalized functions or distributions in meteorology. 1961. (94 blz., 6 fig.) 6,50

80. H. Timmerman. The influence of topography and orography on the precipitation in the Netherlands. 1963. (49 blz., 37 fig., en 5 tab.) 6,50

81. In voorbereiding.

82. In voorbereiding.

83. Fenologisch en faunistisch onderzoek over boomgaardinsekten. Onder red. van H. J. de Fluiter, P. H. van de Pol, en J. P. M. Woudenberg. 1964. (226 blz., 84 fig. en 57 tab.) 9,25

423002*-148