Martin Duke

1514 Walnut St., Berkeley

PRACTICAL
LEAST SQUARES

PRACTICAL LEAST SQUARES

BY

ORA MINER LELAND, B.S., C.E.

Dean of the College of Engineering and Architecture and the School of Chemistry, University of Minnesota. Member of the American Society of Civil Engineers, American Association of Engineers, Society for the Promotion of Engineering Education, American Astronomical Society, etc.

FIRST EDITION
FOURTH IMPRESSION

McGRAW-HILL BOOK COMPANY, INC.
NEW YORK AND LONDON
1921

PRINTED IN THE UNITED STATES OF AMERICA

TO

MY WIFE

WHOSE LOYAL ASSISTANCE
HAS CONTRIBUTED IN GREAT MEASURE
TO ITS PREPARATION,
THIS BOOK IS AFFECTIONATELY
DEDICATED

PREFACE

This book results from the author's experience in teaching the subject of Least Squares and the Adjustment of Observations to classes of civil engineering students at Cornell University. As the time allotted to this work became more and more limited, the available textbooks became less adaptable to the scope of the course. To meet this condition, a series of chapters entitled "Notes on the Adjustment of Observations" was prepared and used as a text. With these notes as a basis, this book has been written.

It is designed particularly for use in short courses of instruction and by engineers and scientists in connection with their private practice. It will not replace the more elaborate treatises on the subject but the author hopes that it will introduce the student directly to the simpler methods of solving the ordinary problems in adjustment.

The plan of the work is essentially practical. After a general introduction devoted to a consideration of the character and occurrence of errors, the adjustment of direct, indirect, and conditioned observations is taken up in detail and illustrated by numerical applications to triangulation, leveling, astronomy, and the derivation of empirical formulas. Not until after this practical treatment of the determination of the best values of the unknown quantities is the precision of observations discussed, together with the computation of the mean square and probable errors of the observations and results. Finally, the principles of probability and the analytical derivation of the Law of Error are given in appendices.

The utility of this arrangement should be obvious. By far the greater number of applications of Least Squares do not require a consideration of the precision of the results nor a knowledge of the mean square or probable errors. Moreover, the subject of the precision is usually the most troublesome part of the work for the student or the beginner to understand. Therefore, the practical methods of adjustment are explained directly and fully, without regard to the probable errors or to the theoretical derivation of the Law of Error. A special effort has been made to explain the procedure in each case as completely as necessary for the beginner as well as the practitioner, even at the risk of criticism for undue length. The usual difficulties experienced by students seem to justify this effort.

In Appendix D there is given an outline of a short course of instruction suitable for civil engineers. This plan was carried out successfully by the author in sixteen lessons. While it is not at all desirable to restrict the work so severely, if no more time can be given to it the course is still very much worth while.

The author is indebted to many excellent works and has endeavored to make specific acknowledgments throughout the book wherever due. In the preparation of the original notes and their application to class instruction, his thanks are due to his former colleagues, Professors P. H. Underwood and L. A. Lawrence, for their assistance and suggestions.

O. M. Leland.

Minneapolis, Minn.
September, 1921.

CONTENTS

CHAPTER III

INDIRECT OBSERVATIONS, OF A FUNCTION OF THE UNKNOWN QUANTITIES

CHAPTER IV

SOLUTION OF NORMAL EQUATIONS

CHAPTER V

OBSERVATIONS OF DEPENDENT QUANTITIES: CONDITIONED OBSERVATIONS

CHAPTER VI

ADJUSTMENT OF TRIANGULATION

CHAPTER VII

EMPIRICAL FORMULAS

CHAPTER VIII

PRECISION OF OBSERVATIONS AND RESULTS AND COMBINATION OF COMPUTED QUANTITIES

COMBINATION OF COMPUTED QUANTITIES

CHAPTER IX

CONCLUSION

APPENDICES

A. HISTORY AND BIBLIOGRAPHY OF LEAST SQUARES

B. PRINCIPLES OF PROBABILITY

C. DERIVATION OF THE LAW OF ERROR

D. OUTLINE OF A SHORT COURSE OF INSTRUCTION

E. TYPICAL CURVES FOR REFERENCE

F. TABLES

PRACTICAL LEAST SQUARES

CHAPTER I

INTRODUCTION

1. Discrepancies among Observations. Measurements made in the field, office, or laboratory directly depend upon readings of scales, circles, micrometers, clocks, watches, etc. The readings may be made to the nearest division, or graduation, or the space between two adjacent graduations may be subdivided by estimation, thus carrying the observation to a greater degree of refinement.[1] When successive settings or pointings of the measuring apparatus are made, upon the same object, the corresponding readings may be the same as the first if the graduations be coarse and the nearest one, only, recorded. But if the divisions be very fine and the readings made with the aid of a magnifier, or reading-glass, and by estimation, there may be considerable variation among them, especially in the last figure which is estimated.

For example, consider the following measurements of a line made with a steel tape in a drizzling rain, using spring-balance, hand-level, and plumb-bobs, the tape being graduated to hundredths.

899.754 ft.	899.763 ft.
.761	.756
.760	.759
.758	.759
.762	.760

If the readings had been made to the nearest hundredth, all after

[1] It is customary to estimate to tenths, although an experienced observer will sometimes record to five one-hundredths when the reading seems to lie between two adjacent tenths, greater than the one and less than the other.

the first would have been alike and 899.76 ft.; if to the tenth only, each reading would have been 899.8 ft., indicating that the care in handling the apparatus would justify the use of a more precise method of making the readings, or that some of the precautions were unnecessary.[1] Thus it will be seen that the observations may be so rough or coarse as to show no variation whatever. Their very agreement, in such a case, might be misleading, as indicating a false precision.

Realizing the occurrence of these small discrepancies among observations, when made with care, the observer makes a number of readings, instead of a single one, and by some method of adjustment adopts a certain value for the observed quantity as a result of his series of observations. If they were made with equal care and under the same conditions, he may consider them to be of equal weight and that none is entitled to preference over the others, in which case it will be reasonable to adopt the simple mean or average of the set as the best value obtainable from these observations. In fact, this adoption of the mean is axiomatic.

2. It will be evident that *absolute correctness* in the observed quantity is *unobtainable* as a result of the observations themselves. In the above example, it would be impossible to determine the length of a line down to a millionth of a foot (the sixth place of decimals), using this method of making the measures. Certainly, then, correctness to an infinite number of places is beyond hope. Moreover, it is impossible to ascertain the correct value of the next figure beyond the limit of our observations. Whatever value may be adopted as a result of adjustment, it should be regarded as but an approximation to the true one, that is, as the best available value within our knowledge.

The discrepancies among the observations of a quantity, then, show that these observations are not quite correct,—that the work is not perfect, in other words, but is attended by errors of observation. The differences between the readings are not the errors themselves but serve to indicate the existence of errors.

[1] Adding a zero after the last observed figure, making the reading, in this example, 899.80 instead of 899.8, is a habit of some beginners which should be studiously avoided.

If it were possible to ascertain the correct value of the observed quantity the *true error* of each observation would be easily found as the difference between the observation and the correct value.[1] But just as it is never possible to know the correct value, so the true errors must be regarded as ideal and indeterminate.

3. Necessity for Adjustment. By making several observations upon a quantity, in succession, two objects are attained, namely, greater *precision* in the resulting mean than in a single observation, and the *check* upon the work afforded by the agreement of the various readings among themselves, within the limits of the small discrepancies above described. The several observations having been made, however, for the purpose of securing a better value of the observed quantity than any one of the separate readings would be likely to be, that is, a value presumably closer to the true or correct value, it is necessary to arrive at, and adopt, some *one* value of the quantity, for use in any computations which may involve it, such use being the probable reason for making the observations in the first place. This necessity arises from the fact that if different values of the same quantity be used in the computation, the results will fail to check.

Similarly, if two or more related quantities, resulting from observations, be used in computations without having been adjusted so as to satisfy the relation between them, the results will be inconsistent and checks upon the computation will be sacrificed. For example, suppose the three horizontal angles of a triangle have been measured in the field and their sum, as usual, fails to equal the theoretical amount, namely, 180° plus the spherical excess of the triangle. In order that the triangle may be computed and the sides checked, the three angles must be adjusted by the application of small corrections so as to satisfy the theoretical sum. Also, if a series of benchmarks be connected

[1] It is well to adopt the rule of subtracting the incorrect or observed quantity from the correct or adjusted one, algebraically, taking account of the signs. The resulting difference, with its sign, is then the correction to be added algebraically to the observed quantity to obtain the adjusted one. Strictly, the *error* has the opposite sign to the *correction*, but the latter is more convenient in most cases, and the use of a fixed rule tends to avoid mistakes. An old expression of this rule is, *Subtract the false from the true.*

by lines of levels, some of which are check-lines forming with the others complete circuits, it is necessary to adjust the differences of elevation so that all of the circuits will close exactly, in order that the difference of elevation between any two benchmarks will be constant when computed through two or more series of lines, that is, by two or more different routes.

Obviously, any computation could be carried out and checked even though the original data were assumed and far from the truth, provided they were not inconsistent. But it is not sufficient that the data be consistent; they must be as near the truth as our knowledge permits if the results are to be of the greatest value. Observations are made for the purpose of securing information with precision, and the results serve as a basis for accurate computations. Therefore, it is important to so combine the observations as to give due consideration to each one and to obtain for each quantity the best value which the given observations can yield, that is, the value which they indicate to be nearest the truth. However, the time and labor involved should not be unreasonable or excessive in view of the objects to be secured.

The process of combining the various observations so as to obtain the best values of the quantities concerned is called the *adjustment of the observations.* The results are referred to as the *adopted, adjusted,* or *corrected* values. The small quantities to be added algebraically to the observations to obtain these adjusted values are known as the *corrections.*

4. Errors of Observation. Every observation made in the process of measurement is likely to be in error from various causes, that is, the actual reading is not the quantity really sought—is not what it would be if conditions were ideal and perfection attainable. Some of these causes are beyond the control of the observer while others depend entirely upon his skill and personality. For example, the altitude of a star is measured with a surveyor's transit. The star appears higher than it really is, owing to atmospheric refraction. The instrument is never in perfect adjustment, so that when the star is seen on the horizontal thread the vertical circle does not show the correct altitude of the line of sight. Moreover, the observer himself may have the habit of noting the time when a star crosses a thread a fraction of a second too late. Then he, or his recorder, may make a mistake of a whole minute in taking the time from his watch. And finally, he reads the vertical circle vernier to the nearest half-minute, perhaps, with a possible error, therefore, of one-fourth of a minute.

It is customary to include the effects of all influences such as those illustrated in the above example in the term, errors, and to classify them as Systematic or Constant Errors, Mistakes or Blunders, and Accidental Errors of Observation.

5. Systematic or Constant Errors occur in accordance with fixed laws or are constant during a set of observations made under unvarying conditions. Their effects are eliminated from observations, as far as our knowledge permits, in two ways: first, by the application of corrections computed from the known laws of the occurrence of the errors; and second, by making the observations according to a prearranged plan so that the conditions will be reversed during half of the set, changing the signs of the corresponding systematic errors; these therefore neutralize those of the other half-set when the observations of the whole set are combined.[1] Systematic errors are divided into three classes, namely, Theoretical, Instrumental, and Personal Errors.

6. Theoretical Errors conform to certain laws from which their effect upon observations made under given conditions may be computed and corresponding corrections applied, as soon as these laws are known. Refraction and aberration of light, expansion of metals with rise of temperature, and dip of the horizon are examples. The form of a law is usually determined theoretically but its constants may result from observations. Theoretical errors are not errors in the sense of being accidents or inaccuracies, but

[1] This arrangement of a program for observing, so as to eliminate systematic errors, is exceedingly important. Observers and computers should always be on the lookout for new and unforeseen sources of these errors, as the observations may not reveal them, and the results, apparently good, may be erroneous to a surprising degree. The experience of the observer is invaluable in his study of the conditions under which his observing is done, with this end in view. As our knowledge of the sources of error increases, so does our ability to bring the results of observations closer to the truth. (See Wright and Hayford: Adjustment of Observations, Art. 201.)

rather are the effects of certain influences which operate to prevent the observer's seeing or reading directly the quantity which he seeks in his observations. They are included in the classification and study of errors merely as a matter of convenience and as a result of custom.

7. Instrumental Errors may be defined as imperfections in the construction or adjustment of instruments, or the effects of those imperfections upon observations made with the instruments. Among these may be mentioned the graduation errors of scales and circles, eccentricity of circles, inequality of pivots, collimation error, and error of runs in a micrometer microscope. They may be determined by measurement and the corresponding corrections applied to the observations, or the observing plan may be such as to eliminate their effects.

8. Personal Errors are generally referred to as *Personal Equation*. They depend upon the habits of the observer and his physical condition. They result, frequently, from the habit of always setting the thread of a telescope slightly to one side of the object sighted, or of always noting the time or giving a signal too early or always too late. No one can hope to be free from such a tendency, and some of the best observers the world has ever known have had unusually large personal equations. Good, steady observers in normal physical condition will have nearly constant personal equations, whether large or small, and this steadiness of habit is more important than that the error be small in amount. If the observations be differential in character, the personal equation of the observer may have no effect, if it be constant and if all the readings be made by him. This, for example, would be the case in leveling, if the rod-target were always placed too high or always too low and by the same amount. Similarly, it may not affect the measurement of angles in triangulation. But if different observers be involved, the results may be affected by the sum or difference of their personal equations.

The effect of this error may be eliminated, in some cases, by an exchange of observers, as in telegraphic longitude determinations; or, its amount may be determined by special experiments or apparatus, for each observer, then *assumed to be constant* and applied as a correction to his subsequent observations of the same kind when made under the same conditions, especially as regards his personal comfort and health. However, the personal equation of an observer must not be assumed as constant for any great length of time, and there is always danger in assuming it constant at all. It is safer to determine it at different times and to interpolate for its value between these results. Depending upon personal peculiarities, it follows no law and is often the most troublesome source of error to which observations are subject. Fortunately, it is small in amount in most cases.

9. Mistakes or Blunders are irregular in their occurrence, obeying no law, and are relatively large in size. They result from haste and carelessness, frequently, on the part of the observer, during temporary lapses, perhaps, from his customary vigilance. He may call out to his recorder one number while reading and thinking another; he may read the wrong division of a circle or scale; or he may read a clock wrong by a whole minute while he is estimating tenths of a second. He may turn the wrong tangent-screw while repeating angles, the rod-clamp may slip during leveling, or the wrong object may be sighted in triangulation or azimuth work. The remedy lies in uninterrupted care on the part of the observer to avoid these blunders, and watchfulness by the recorder to detect them in any inconsistencies among the readings. Herein lies one of the chief virtues of a good recorder.

10. Accidental Errors of Observation, or simply Accidental Errors, is a name given to a specific class of errors in connection with the adjustment of observations by the Method of Least Squares. They are purely errors of observation and have no relation to systematic errors or the large mistakes already described. They are small, for the most part, and their presence is indicated by the discrepancies among a series of readings upon a fixed object which have been made with the utmost care and precision, with an instrument which can be read to a greater degree of refinement than the pointings can be made by the observer. These errors are never known exactly because the true or correct value of the quantity observed is never known, as has been explained in a previous article. Thus it is stated that they are *indicated* by the discrepancies, not that they *are* the discrepancies themselves.

11. For example, suppose readings are made by a skillful observer using a micrometer microscope, upon a graduation line or scratch of a standard meter bar, the whole being enclosed in a vault of constant temperature so that conditions are steady. Further, suppose the observer to be able to set the parallel threads so as to be equidistant from the scratch within 10 microns [1] and that the micrometer reads directly to five microns and by estimation to one half-micron, that is to 0.0005 millimeter. The readings, then, might run as follows, the unit being one division of the micrometer head (equal to 0.005 mm.):

d	d
46.4	45.9
45.6	45.3
46.0	46.1
46.1	45.8
45.9	45.8
46.6	45.2
46.7	46.1
45.4	46.8
46.5	45.1
45.9	46.3

By assumption, the conditions are very favorable for precise work and the observer is skillful and is using great care in making the readings; nevertheless, there are discrepancies and the readings have a total range of 1.7 divisions. These are the discrepancies which indicate the presence of accidental errors of observation. They are so small as to be beyond the control of the observer, as he is assumed to make each separate pointing as carefully as he can. It may be noted, also, that most of the discrepancy is due to the errors of pointing, that is, setting the threads on the mark, as the estimation of tenths of a division of the head would seldom be in error by a whole tenth.

12. In other examples, the discrepancies might be made up of accidental errors of several different kinds, such as pointing the

[1] A micron is one one-thousandth of a millimeter or one one-millionth of a meter. It is the unit used in very precise measurements of length by means of micrometer microscopes.

telescope upon a signal, setting the threads of the microscope upon a division of the circle, and reading the micrometer head. Other sources of errors which may have the nature of accidental errors are the unsteadiness of the atmosphere and that of instrument supports, and rapid changes of temperature. However, the foregoing example of simple, direct readings is a clear illustration of the occurrence of accidental errors without complication. If the micrometer head had been graduated directly into one thousand parts instead of one hundred, to be read with a magnifier, the error of estimation in reading it would have disappeared and the discrepancies might have been ascribed entirely to the accidental errors of setting the threads upon the division on the circle,—the simplest kind of a case.

13. Accidental Errors, only, Considered in Adjustments. It has been shown that the effects of systematic errors are eliminated by corrections or by the observing program, as far as they are known to exist; and that the mistakes, or blunders, are avoided by the exercise of care and vigilance, as much as possible. Of all the kinds of errors, then, there remain the accidental ones, still affecting the observations, and it is to minimize the effects of these errors that adjustments are made. In all that follows in this work, therefore, only this special class, the accidental errors, will be considered, except as others may be specifically mentioned.

14. Assumption of the Arithmetic Mean.[1] When each observation or reading has been made with the same care and under the same conditions as all the others of a set made upon a certain quantity, so that all are of equal value, or weight, there is no reason for preferring any one to any other; the mean, or average, of them all must be regarded, then, as the best value of the observed quantity which can be obtained from the given set of observations. The soundness of this principle is so evident that it is adopted as the fundamental assumption in developing the theory of the adjustment of observations. The mean should be regarded, not as the true value of the observed quantity, but rather as the

[1] The word *mean*, in this work, is understood to refer to the *arithmetic* mean, or average, in every case. The *geometric* mean is the square root of the product of two quantities.

nearest approximation to it that the given observations will yield, and subject to improvement if other or better observations should become available.

15. Residuals (v). The difference between an observed value of a quantity and the adopted one is known as the residual of that observation. It should be taken in the sense, *adopted minus observed,* for consistency in sign. If the adopted value, the mean, for example, be the nearest approximation to the truth, then the residuals obtained with that value would be the nearest approximations to the actual or true errors of the observations, to the extent of our knowledge. The occurrence and behavior of the residuals, then, will be our best indication as to the occurrence of the true errors. In fact, we may reasonably assume that the errors and the residuals conform to the same laws. In the investigation of such laws, therefore, it may be convenient, sometimes, to use the terms somewhat indiscriminately,—to use the word error when residual is intended.

16. Regularity in the Occurrence of Accidental Errors. At first thought, it may seem strange that there should be any method at all in the occurrence of errors which are so small and so evidently the result of accident or inaccuracy. However, it has been found from a large number of investigations of observations of almost every conceivable sort, that these errors occur not only with regularity but in conformity to a definite law, of which the general form is the same for all kinds of observations. This law of the occurrence of errors, or Law of Error, as it is called, is expressed in the form of an equation which has been completely derived, and tested, later, in a multitude of cases, with entire satisfaction. In accordance with this law of error, the Method of Least Squares has been devised and demonstrated for the adjustment of observations.

17. Curve of Error. As an example, let us consider a large set of direct observations, say 500 of them, such as the readings of the micrometer microscope in the example of Art. 11, page 8. By taking the mean of the entire series and subtracting from it each separate reading, the residuals are obtained. Counting the residuals of each size and sign, we find that there are 36 of $+0.1$, 35 of -0.1, 33 of $+0.2$, 34 of -0.2, etc., the sum of all the numbers being, of course, 500. These results are plotted as rectangular coordinates, the magnitude of the error on the horizontal axis, plus on the right and minus on the left of the origin, and the corresponding number of errors of that size on the vertical axis, upward. Thus one point is plotted for each size of error, and for each sign. A smooth curve is then drawn so as to follow the points as closely as possible, with the result shown in Fig. 1:

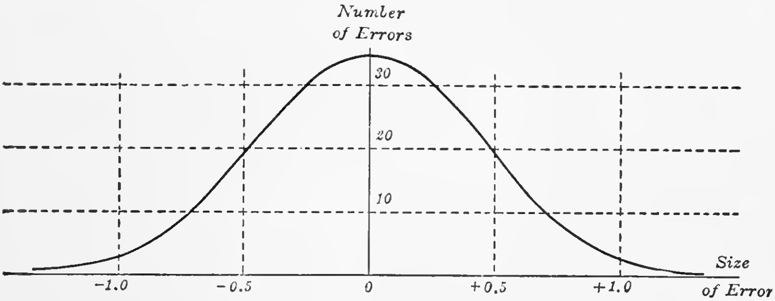


FIG. 1. Occurrence of Errors or Residuals

The form of the curve is typical of all those constructed from observations in this manner, and it is called the Curve of Error, or the Curve of Probability of Error, since an ordinate to the curve may represent the probability[1] of the occurrence of an error as well as the number of times that error occurs.

18. Assumptions as to the Occurrence of Errors. The error curve has three properties which are evident from inspection: first, it is symmetrical about the vertical axis; second, it has a maximum point where it crosses that axis; and third, it approaches the horizontal axis so gradually as to appear asymptotic. Generalizing from these properties, the following assumptions, or axioms, are obtained, as to the occurrence of errors in any large set of observations:

1. Positive and negative errors of the same magnitude occur with equal frequency; they are equally probable.

2. Small errors occur more frequently, or are more probable, than large ones.

[1] The probability of an event is directly proportional to the number of times it occurs. See Appendix B, for Principles of Probability.

3. Very large errors seldom occur; they are likely to belong in the class of mistakes rather than that of accidental errors.

It should be remembered that the number of observations in a set is assumed to be large. The smaller the set, the less closely will the residuals conform to the ideal conditions, such as that of the first of these assumptions, but even in a small set they will approximate to their ideal occurrence. Obviously, the larger the number of observations, the more closely should the mean approach the true value of the quantity observed, in so far as the accidental errors are concerned.

19. Law of Error. The general equation of the error curve, or curve of probability, may be derived [1] from the assumptions of the last article together with the principle of the mean (Art. 14, page 9). The curve is seen to be continuous, and it is of special importance to note that the number of errors, or the probability of an error, is a function of the size of the error. The algebraic principles of probability, also, are involved in the derivation. The resulting Law of Probability of Error may be stated thus:

$$p = \phi(\Delta) = \frac{h}{\sqrt{\pi}} e^{-h^2\Delta^2} \qquad (1)$$

in which p is the probability of an error Δ in a set of observations for which h is a computed constant; e is the base of natural logarithms; and $\pi = 3.1416^+$. The constant h has the value $\frac{1}{\epsilon\sqrt{2}}$, in which ϵ is a constant for each separate set of observations,[2] and serves to change the general equation into a specific one for the particular set of observations under consideration.

20. Tests of the Law of Error. The law may be tested by applying it to many different kinds of observations so as to ascertain whether the residuals occur in conformity with it. Conversely, if the law be accepted as applicable to all observations,

[1] This derivation may be found in Appendix C.

[2] ϵ will be defined farther on as the mean square error of a single observation. Its value, for a given set of observations, depends upon their precision, and is determined from the residuals.

the quality of a given set could be tested by the same method. In general, then, it is a process of comparing theoretical results with observed ones, or theory with practice. The method consists of the comparison of the number of residuals, in the given series, which lie between certain limits, with the number of errors which ought to lie between those limits according to the Law of Error. For example, ϵ having been computed for the given observations, the probability of an error between 0.00 and 0.30, say, is determined by integration and substitution in the equation (1).[1] (It will always be less than unity, from the principles of probability.) Multiplying the total number of observations in the set by this probability gives the number of errors which ought to lie between the assumed limits according to the law. The residuals which actually lie between those limits may then be counted and their number compared with that obtained from the formula.

Crandall gives an example of a small set of 18 observations of an angle, with the following results, ϵ being 1.66″.

	Number of Errors Less Than				
	0.5″	1″	2″	3″	4″
From theory	4.3	8.1	13.8	16.8	17.8
By actual count	6	8	14	17	17

This agreement is very satisfactory but might be much closer in a larger set of observations.

21. Method of Least Squares. The most probable value of the observed quantity obtainable from a given set of observations will be the one corresponding to the most probable set of errors or of residuals. Consider a set of n observations of equal precision, of which the most probable errors are $\Delta_1, \Delta_2, \Delta_3, \ldots \Delta_n$, respectively. Since the probability of the simultaneous occurrence of several events in a series is the product of their separate probabilities,[2] and the probability of an error, Δ, is, from (1):

[1] The use of the equation is facilitated by its transformation into series from which tables have been computed. See Appendices C and F.

[2] See Appendix B.

$$p=\frac{h}{\sqrt{\pi}}e^{-h^2\Delta^2}$$

it follows that the probability of the simultaneous occurrence of the errors $\Delta_1, \Delta_2, \Delta_3, \ldots \Delta_n$, will be

$$P=\left(\frac{h}{\sqrt{\pi}}e^{-h^2\Delta_1^2}\right)\left(\frac{h}{\sqrt{\pi}}e^{-h^2\Delta_2^2}\right)\ \ldots\ \left(\frac{h}{\sqrt{\pi}}e^{-h^2\Delta_n^2}\right)$$

that is

$$P=\left(\frac{h}{\sqrt{\pi}}\right)^n e^{-h^2(\Delta_1^2+\Delta_2^2+\ \ldots\ +\Delta_n^2)} \tag{2}$$

But since these errors are to be the most probable ones, P must have its maximum value. As h, π, n, and e are constant in a given problem, and the exponent of e is always negative, the expression will be a maximum when the exponent of e is a maximum, algebraically, that is, when the sum

$$\Delta_1^2+\Delta_2^2+\Delta_3^2+\ \ldots\ +\Delta_n^2 \text{ is a minimum} \tag{3}$$

Thus, the most probable value of the observed quantity, or the *best value*, in other words, obtainable from the given set of observations, will be the one for which the sum of the squares of the errors, or of the residuals, likewise, is a minimum. This is called the Principle of Least Squares and the method which is based upon it, for the adjustment of observations, is known as the Method of Least Squares. It was first published by Legendre, in 1806, although used by Gauss as early as 1794.[1] In the general case, involving the determination of several quantities, and observations of unequal weight, it provides that the most probable values of the unknown quantities will be those for which the sum of the weighted squares of the residuals is a minimum. This form will be discussed later (Art. 34).

22. Number of Observations. In the development of the Method of Least Squares, it is assumed that the number of observations is large. The assumptions as to the occurrence of errors approach the truth more closely as the number of the errors increases. However, if the method be applied to small sets of

[1] See Appendix A.

observations, although the results may be farther from the correct values, still they may be regarded as the best ones obtainable from the given observations, which is sufficient warrant for the use of the method under such unfavorable conditions. It is unreasonable to generalize too greatly from a very small set of residuals, as to the precision of a result, but it is still permissible to take the mean of even two observations, if they be the only available data.

Regarding the number of observations, it must be remembered that *no adjustment is possible unless there are more observations than unknown quantities.* If the number be less, the unknowns cannot be determined without additional information or assumption. If the number be equal to that of the unknowns, there is only one solution, namely, the rigid, algebraic one by means of simultaneous equations.

23. Two Uses of Least Squares. The Method of Least Squares is essentially a practical subject, being devoted to the solution of numerical problems. Its applications may be divided into two classes: first, the determination of the *best values* of the unknown quantities obtainable from given observations, that is, the adjustment of observations; and second, the investigation of the *precision* of the observations and the results, and the influence of errors upon them. These two uses of the method are quite independent; most problems require adjustment, but the precision may not be investigated at all unless the results are to be compared with those of other observations. In this treatment of the subject, therefore, immediate attention will be given to the adjustment of the various kinds of observations, but the determination of the precision of the results will be postponed to a later chapter.[1]

24. Classification of Problems. In the following chapters, the typical problems are such as the engineer frequently encounters in field work. A certain method of solution is adopted for each type. The adjustment of the three great classes of observations is taken up in the usual order, namely:

[1] Chapter VIII.

Direct Observations of One Quantity,

Indirect Observations, of a Function of the Unknown Quantities, and

Observations of Conditioned Quantities.

Following these, the investigation of precision and the propagation of error will be explained. It is important that the student become familiar with the characteristics of these classes of problems and with the method of solution of each type. The special problem of the derivation of empirical formulas and constants will be treated in a separate chapter.[1]

[1] Chapter VII.

CHAPTER II

DIRECT OBSERVATIONS OF ONE QUANTITY

25. Direct Observations: Readings. In their simplest form, direct observations consist of single readings made upon various kinds of apparatus used in measurements, such as scales, circles, micrometers, and timepieces. The example in Art. 11, of micrometer readings upon a fixed scale, is typical of this class. The conditions under which the readings are made are assumed to be constant or to vary according to a known law so that the discrepancies among the readings may be reduced to the accidental errors of pointing or setting the instrument and of reading.

Usually, however, the conditions are more complex and involve several sources of error. In the example just cited, the temperature may vary, causing the position of the division line on the scale to change. Then if the temperature be read from a mercurial thermometer simultaneously with the micrometer readings, two corresponding sets of direct readings are obtained. Also, when the altitude of a star is observed for time and azimuth, each pointing on the star may be attended by readings of the watch and the horizontal and vertical circles, so that three simultaneous sets of direct readings result.

26. Observations Resulting from a Combination of Readings. It frequently happens, on the other hand, that the so-called observed quantity is the result of two or more separate readings of the same kind. For example, in the measurement of angles by repetition, a single observation is obtained by subtracting the initial reading from the final one and dividing the difference by the number of repetitions, in the case of the direct measure of the angle itself and also, of the reversed measure of its explement, the mean of the two results being taken.[1] In the measurement of a base line, each observed length is the sum of several tape-lengths; the elemental observations consisting of placing the rear scratch on the tape in contact with a scratch on a marking-plate and of making a mark on a plate opposite the scratch at the forward end of the tape. Similarly, the observed difference of elevation between two benchmarks consists of the algebraic sum of a series of fore- and back-sight readings of the rod. It is customary, in all such cases, to consider the result of a single complete measurement to be the observed quantity, even though it consist of a combination of separate readings. In its general sense, therefore, the term, direct observation, may be taken to mean a single measurement of the quantity desired.

27. The Mean. The adjustment of direct observations of a single quantity consists in taking their mean as the best, or most probable, value obtainable from the given observations, as explained in Art. 14. That this is in accordance with the principle of least squares, may be shown as follows:

Let M_1, M_2 . . . M_n represent a series of observed values; x_0, the best value of the observed quantity; and v_1, v_2, . . . v_n, the corresponding residuals, n being the number of observations. Then, for each observation there results an observation equation, thus:

$$\begin{aligned} x_0 - M_1 &= v_1 \\ x_0 - M_2 &= v_2 \\ . \ . \ . \ &. \ . \ . \ . \\ x_0 - M_n &= v_n \end{aligned} \tag{4}$$

Squaring both members of each equation and adding the resulting equations, we obtain,[2]

$$(x_0 - M_1)^2 + (x_0 - M_2)^2 + . \ . \ . + (x_0 - M_n)^2 = [v^2] \tag{5}$$

[1] A simpler method is to subtract from the reading on the right-hand object the mean of the two readings on the left-hand object (first and last readings) and divide the difference by the number of repetitions. The result is the same.

[2] The square brackets, [ ], indicate the sum of all such terms as are included by them. Thus, $[v^2]$ represents the sum of the squares of all the v's, that is, $v_1^2 + v_2^2 + v_3^2 + \ . \ . \ . + v_n^2$. The symbol, Σ, may be used to indicate summation in the same manner.

According to the principle of least squares, the sum of the squares of the residuals, that is, $[v^2]$, is to be a minimum. Therefore, we differentiate the left-hand member and place the first derivative equal to zero; whence, after dividing by two, we have:

$$(x_0-M_1)+(x_0-M_2)+\ .\ .\ .\ +(x_0-M_n)=0 \tag{6}$$

$$nx_0-(M_1+M_2+\ .\ .\ .\ +M_n)=0,$$

and
$$x_0=\frac{[M]}{n} \tag{7}$$

That is, the best value of the observed quantity—the one for which the sum of the squares of the residuals is a minimum, is the *mean*.

28. Computation of the Mean. Owing to the close agreement of the observations of which the mean is to be taken, it is possible, often, to abridge the numerical work by the assumption of an approximate value of the mean. Suppose the mean of the following 16 quantities to be desired

	$+v$	$-v$		$+v$	$-v$	
1463.49768	....	4	1463.49764	0	....	Assumed approx.
1463.49754	10	....	58	6	....	value, 60
1463.49763	1	....	66	....	2	
1463.49765	....	1	63	1		
1463.49759	5	....	65	....	1	Sum, +65
1463.49771	....	7	60	4		
1463.49767	....	3	67	....	3	Mean, +4
1463.49765	....	1	70	....	6	Remainder, +1
	16	16		11	12	

Mean, 1463.49764 $\qquad [v]=-1$

In the first place, it is evident from inspection that all but the last two figures are the same in all the quantities, so that there is no need of writing them repeatedly. It is sufficient to write the first number in full and thereafter only the last two figures. Sometimes, in a long column of numbers, the constant part will change suddenly to another one, in which case it is well to write the last as well as the first of each series in full.

Also, by inspection, it will be seen that the next to the last figure in the mean will probably be 6. Thus we may take 60 as an approximate value of the last two places, calling 54, for example, -6, and 71, $+11$. Then, adding mentally the figures in the last place with 60 as a basis, we obtain the sum, $+65$, and the mean, $+4$, so that the full mean is 1463.49760+4 in the last place, or 1463.49764. This process may be simplified still more by combining a 5 and a 7 in the next to the last place, as their mean is 6, without modifying their last figures. Thus in the above example, 59 and 71 would be added directly as 10 instead of -1 and $+11$.[1]

29. Control or Check of the Mean. Substituting equations (4) in (6) of Art. 27,

$$v_1+v_2+v_3+\ldots+v_n=0 \qquad \text{or,} \qquad [v]=0 \tag{8}$$

That is, *the sum of the residuals should be zero,* or the sum of the positive residuals should be equal to that of the negative ones. This check is very important and should be used whenever practicable. It will be satisfied rigidly unless there is a remainder when the sum of the observations is divided by their number to obtain the mean. In this case, the check fails by just the amount of the remainder but with the opposite sign, so that the mean is verified, nevertheless. In the example in the preceding article, the sum of the residuals is -1, and the remainder in taking the mean is $+1$, so that the mean was correctly computed.

30. Weighted Observations. Thus far, we have considered only observations of equal quality or precision. In the general case, however, one observation of a series may be better than another, for some reason, and entitled to have a greater influence upon the result. When all of the observations of a set are not of the same quality or worth, they are called weighted observations, or are said to be of unequal weight.

31. Definition of Weight (w). By the weight of an observation is meant its relative value among the others of a set. It is

[1] The beginner will do well to learn to add mentally by combinations of two or three figures at once, particularly those whose sum is 10, as 6 and 4, 7 and 3, or 5, 2, and 3, even though another figure intervenes.

expressed as a number, and being strictly relative, may be multiplied by any factor so long as all the others in the set are multiplied by the same quantity. Thus, the weights may be integral or fractional. If one observation has a weight of 3 and another a weight of unity, the first may be considered as the mean of three observations of the same size, each of which has the weight unity. The weights could be stated as 6 and 2, as 1 and $\frac{1}{3}$, or as 0.153 and 0.051, as well as 3 and 1.

32. Sources of Weights. Either the observer or the computer may assign the weights to the observations and it is largely a matter of judgment. If the observer assigns them, during the observing, he has the right to do it by *estimation* or arbitrarily. For instance, in the measurement of angles in triangulation, the atmosphere may be so unsteady during one observation that he will give to that particular result a weight of one-half that of the others. Or he may note in his record the fact that the atmosphere was very unsteady at that time, and leave to the computer, in the office or at headquarters, the duty of assigning a low weight to that observation, when making the adjustment. Of course, the computer might give it a weight of 0.8 instead of 0.5, and thus change the result somewhat. Or, an arbitrary rule might be agreed upon so that both would assign the same weight under the same circumstances. Similarly, two benchmarks may be connected by two lines of levels giving discordant results. If one run were made during a high wind or with a careless rodman, it might be given a lower weight than the other.

In the second place, weights may be assigned upon the *number of observations*, as a basis. If one measurement of an angle be made with three repetitions and another with six, the second may be given twice the weight of the first.

Finally,[1] the assignment of weights may be governed by *theory*. In the determination of time by transits of stars across the meridian, the motion of a star near the equator will be more rapid than that of one of greater declination, and the rapidly moving one can be observed more accurately than the other. Therefore, a system

[1] For the determination of weights from mean square errors, see Art. 156, Chapter VIII, Combination of Computed Quantities.

of weights has been devised which depends upon the declinations of the stars.

33. The Weighted Mean. The best value of the observed quantity which is obtainable from a given series of weighted observations is known as the Weighted Mean. To determine it, each observation is multiplied by its weight and the sum of these products is divided by the sum of the weights. The analogy of this process to the determination of the simple mean will be evident from an example.

Let it be required to adjust the following set of four weighted observations of an angle, the weights, w, being determined from the number of repetitions and the notes as to weather conditions:

M	w			$+v$	$-v$	wM	$+wv$	$-wv$
73° 18′ 42.16″	3	2.16 or	+.16	...	9			
41.96	2	2.16	+.16	...	9			
41.70	2	2.16	+.16	...	9	+.48	...	27
42.23	4	1.96	−.04	11				
		1.96	−.04	11	...	−.08	22	
	11	1.70	−.30	37				
Use 42.00 as		1.70	−.30	37	...	−.60	74	
Approx. value		2.23	+.23	...	16			
		2.23	+.23	...	16			
$\frac{+0.72}{11} = +0.07$		2.23	+.23	...	16			
		2.23	+.23	...	16	+.92	...	64
Mean, 42.07″		22.72	+.72	96	91	+.72	96	91

By writing each observation a number of times equal to its weight, and by using 42.00″ as an assumed or approximate value of the mean, the third column is obtained. According to the definition in Art. 31, this reduces all the quantities in these columns to the same, unit weight, and their number is the sum of the weights. Therefore, their mean is the best value, and by the methods of Art. 28, this is 42.07″, with residuals shown in the columns headed $+v$ and $-v$. The mean is checked by its remainder, −5, against the sum of the residuals, +5.

It is evident that instead of writing the first observation three times in the third and fourth columns, it will be easier to multiply it by three and write the product, and similarly with the other observations and their weights; the sums would be unchanged. Likewise, the residuals may be multiplied by the weights of the corresponding observations and the products noted instead of writing all the separate residuals. Thus, the last three columns are obtained, giving the same results as the preceding ones. The following *rule*, therefore, is given for the adjustment of direct observations of unequal weight, whether the weights be integral or fractional: Multiply each observation by its weight and divide the sum of the products by the sum of the weights, to obtain the weighted mean; and multiply each residual by the weight of the corresponding observation, adding the products algebraically to obtain the sum of the weighted residuals.

34. Principle of Least Squares for Weighted Observations. Let $M_1, M_2, M_3, \ldots M_n$ represent a set of n observations having the respective weights, $w_1, w_2, w_3, \ldots w_n$, and let x_0 be the best value of the observed quantity, with $v_1, v_2, v_3, \ldots v_n$ as the corresponding residuals.[1] Considering each observation of weight w to be the mean of w equal observations of weight unity, the residual of each of these latter observations would be the same as that of the original one, but there would be w of them. As stated in Art. 21, for the best value of the observed quantity, in the case of equal weights, the sum of the squares of the residuals will be a minimum. Therefore, to express this minimum for weighted observations, each residual must be written a number of times, w, equal to the weight of its observation. Thus,

$$(v_1^2+v_1^2+v_1^2+\ldots \text{ to } w_1 \text{ terms})+(v_2^2+v_2^2+v_2^2+\ldots \text{ to } w_2 \text{ terms})$$
$$+\ldots+(v_n^2+v_n^2+v_n^2+\ldots \text{ to } w_n \text{ terms}) \text{ is to be a minimum;}$$

that is,

$$w_1v_1^2+w_2v_2^2+\ldots+w_nv_n^2 \text{ must be a minimum} \tag{9}$$

or the sum of the weighted squares of the residuals must be a minimum. Substituting for each v in (9) its value, x_0-M, with the corresponding subscripts,

$$w_1(x_0-M_1)^2+w_2(x_0-M_2)^2+\ldots+w_n(x_0-M_n)^2 \text{ is to be a minimum.}$$

[1] Reference to the numerical example of the preceding article will be of assistance in following these steps.

Differentiating this expression and placing the first derivative equal to zero, for the minimum, we have, after canceling the factor 2:

$$w_1(x_0-M_1)+w_2(x_0-M_2)+\ . \ . \ . \ w_n(x_0-M_n)=0 \qquad (10)$$

Combining terms,

$$x_0(w_1+w_2+w_3+\ . \ . \ . \ +w_n)$$
$$-(w_1M_1+w_2M_2+w_3M_3+\ . \ . \ . \ +w_nM_n)=0$$

and

$$x_0=\frac{[wM]}{[w]} \qquad (11)$$

that is, the best value of the observed quantity, for which the sum of the weighted squares of the residuals is a minimum, is the weighted mean, obtained by multiplying each observation by its weight and dividing the sum of the products by the sum of the weights.

35. Control or Check of the Weighted Mean. If in (10), above, v be substituted for x_0-M, we have

$$w_1v_1+w_2v_2+w_3v_3+\ . \ . \ . \ +w_nv_n=0 \qquad (12)$$

or, *the sum of the weighted residuals should equal zero.* As was the case, however, in the control of the simple mean, the actual sum of the weighted residuals should equal the remainder obtained with the weighted mean but with the opposite sign. This is illustrated in the example of Art. 33.

36. Weighted Mean of Two Quantities. The solution of this special case is particularly convenient and instructive. With the usual notation, let M_1 and M_2 be the two given quantities, whose weights are w_1 and w_2 respectively, and let x_0 be their weighted mean. Then from (11),

$$x_0=\frac{w_1M_1+w_2M_2}{w_1+w_2} \qquad (13)$$

Adding and subtracting w_2M_1 from the numerator,

$$x_0=\frac{w_1M_1+w_2M_2+w_2M_1-w_2M_1}{w_1+w_2}$$

$$=\frac{M_1(w_1+w_2)+w_2(M_2-M_1)}{w_1+w_2}$$

$$=M_1+\frac{w_2}{w_1+w_2}(M_2-M_1) \qquad (14)$$

Also, owing to the symmetry of (13), the subscripts may be interchanged, and therefore,

$$x_0 = M_2 + \frac{w_1}{w_1 + w_2}(M_1 - M_2)$$

Thus, the weighted mean may be found by correcting one of the quantities by an amount equal to the difference between the two quantities multiplied by the weight of the *other* and divided by the sum of the weights. Obviously, the mean lies between the two quantities, so the sign of the correction will be evident. The weighted mean divides the interval between the two quantities in the inverse ratio of the weights of the adjacent quantities.

For example, the weighted mean of

6.784 Wt. 7

and 6.743 Wt. 2 is $6.784 - \frac{2}{9} \times 41 = 6.784 - 9 = 6.775$

the unit, for the correction, being conveniently taken in the last decimal place. Similarly, the correction to the second quantity would be $+\frac{7}{9} \times 41$, with the same result.

CHAPTER III

INDIRECT OBSERVATIONS, OF A FUNCTION OF THE UNKNOWN QUANTITIES

37. Indirect Observations are those in which the observed quantity is related to the desired unknown quantities through a known formula or function. The observed quantity is expressed as an explicit function of the unknowns, which are usually two or more in number, and is, therefore, the observed value of the function. It may be that the unknowns cannot be separated so as to be observed directly, and that they can only be determined in combination. They are assumed to be mutually independent; each may vary without causing a corresponding variation in the others. Moreover, the number of observations must be greater than that of the unknown quantities, as stated in Art. 22.

38. The General Function may be algebraic, logarithmic, exponential, or trigonometric, and simple or complicated. However, it is always possible to reduce such a general function to the linear form, that is, to the first degree, either by taking the logarithm of each member or by developing the function by Taylor's Theorem and neglecting the squares, products, and higher powers of the small increments involved.[1] Furthermore, the great majority of problems are concerned with the simplest form of function, namely, the algebraic one of the first degree. Therefore, we shall here consider only this linear form.

39. The Linear Function between the unknowns, x, y, z, etc., will have the following general form,

$$ax+by+cz+\ldots+k \tag{15}$$

in which a, b, c, etc., are known numerical coefficients or factors and k is the constant term. As usual, the signs represent algebraic addition and the quantities may be positive or negative.

40. Observation Equations are the algebraic statements of the separate observations. Thus, if M_1, M_2, ... M_n be the

[1] See Arts. 119–121.

observed values of the function, with the respective weights, $w_1, w_2, \ldots w_n$, the observation equations would be,

$$\begin{array}{ll} a_1x+b_1y+\ldots+k_1=M_1 & \text{Wt. } w_1 \\ a_2x+b_2y+\ldots+k_2=M_2 & \quad w_2 \\ \cdots\cdots\cdots\cdots\cdots & \\ a_nx+b_ny+\ldots+k_n=M_n & \quad w_n \end{array} \tag{16}$$

Their number is the same as that of the observations, and each subscript indicates its equation and observation.

If X, Y, Z, etc., be the best or most probable values of x, y, z, etc., to be obtained from the given observations, the substitution of these values in the above equations (16) would show a residual, v, for each equation, inasmuch as the observations are subject to error and no set of values of the unknowns would be likely to satisfy exactly any one of the observation equations. The ideal form of these equations, therefore, would be,

$$\begin{array}{ll} a_1X+b_1Y+c_1Z+\ldots+k_1=M_1+v_1 & \text{Wt. } w_1 \\ a_2X+b_2Y+c_2Z+\ldots+k_2=M_2+v_2 & \quad w_2 \\ \cdots\cdots\cdots\cdots\cdots & \\ a_nX+b_nY+c_nZ+\ldots+k_n=M_n+v_n & \quad w_n \end{array} \tag{17}$$

Transposing the M of each equation, and representing the difference, $k-M$, of the two constant quantities by the constant term, l, we have for the observation equations,

$$\begin{array}{ll} a_1X+b_1Y+c_1Z+\ldots+l_1=v_1 & \text{Wt. } w_1 \\ a_2X+b_2Y+c_2Z+\ldots+l_2=v_2 & \quad w_2 \\ \cdots\cdots\cdots\cdots\cdots & \\ a_nX+b_nY+c_nZ+\ldots+l_n=v_n & \quad w_n \end{array} \tag{18}$$

which are sometimes called *Residual Equations.*

41. Adjustment of Indirect Observations of Unequal Weight. For the best values of the unknown quantities, the sum of the weighted squares of the residuals is to be a minimum. That is,

$$w_1v_1^2+w_2v_2^2+w_3v_3^2+\ldots+w_nv_n^2 \text{ must be a minimum} \tag{9}$$

Since x, y, z, etc., are independent of one another, it follows that the first derivative of the above expression with respect to each of them must separately equal zero, for the minimum. Differentiating (9), therefore, and canceling the factor, 2, from each equation, we have:

$$
\left.\begin{aligned}
&w_1 v_1 \frac{dv_1}{dX} + w_2 v_2 \frac{dv_2}{dX} + \ldots + w_n v_n \frac{dv_n}{dX} = 0 \\
&w_1 v_1 \frac{dv_1}{dY} + w_2 v_2 \frac{dv_2}{dY} + \ldots + w_n v_n \frac{dv_n}{dY} = 0 \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots
\end{aligned}\right\} \tag{19}
$$

There will be one equation for each of the unknown quantities. The differential coefficients in the first of these equations are the coefficients of X in the successive equations (18), those in the second are the coefficients of Y, etc. Substituting the value of the v's from (18), in the equations (19), then, we obtain

$$
\left.\begin{aligned}
&w_1 a_1(a_1X + b_1Y + \ldots + l_1) + w_2 a_2(a_2X + b_2Y + \ldots + l_2) \\
&\qquad + \ldots + w_n a_n(a_nX + b_nY + \ldots + l_n) = 0 \\
&w_1 b_1(a_1X + b_1Y + \ldots + l_1) + w_2 b_2(a_2X + b_2Y + \ldots + l_2) \\
&\qquad + \ldots + w_n b_n(a_nX + b_nY + \ldots + l_n) = 0 \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots
\end{aligned}\right\} \tag{20}
$$

Whence carrying out the products indicated, and adding the similar terms,

$$
\left.\begin{aligned}
&[wa^2]X + [wab]Y + [wac]Z + \ldots + [wal] = 0 \\
&[wba]X + [wb^2]Y + [wbc]Z + \ldots + [wbl] = 0 \\
&[wca]X + [wcb]Y + [wc^2]Z + \ldots + [wcl] = 0 \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots
\end{aligned}\right\} \tag{21}
$$

These are called the *Normal Equations*, as they are the same in number as the unknown quantities, and, therefore, may be solved simultaneously to determine the latter. It will be seen in the equations (20) that the first normal equation is formed by multiplying the left-hand member of each observation equation by its weight and the coefficient of X in that equation, and adding all the resulting products. Likewise, the second normal equation is formed by multiplying each observation equation by its weight and the coefficient of Y, and adding the products, and so on through the series of unknown quantities.

The adjustment, then, consists in forming from the given observation equations a set of normal equations, the same in number as the unknown quantities, the solution of which as simultaneous equations will give the best values of those quantities.

42. Observations of Equal Weight. This is a special case of the foregoing, in which each weight may be replaced by unity so that the w's disappear from the normal equations (21), resulting, therefore, in the following simpler form:

$$\begin{aligned}
&[a^2]X+[ab]Y+[ac]Z+\ \ldots\ +[al]=0 \\
&[ba]X+[b^2]Y+[bc]Z+\ \ldots\ +[bl]=0 \\
&[ca]X+[cb]Y+[c^2]Z+\ \ldots\ +[cl]=0 \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots
\end{aligned} \tag{22}$$

For purposes of illustration, it will be convenient to use these equations (22) rather than the longer ones in which the weights are included.

43. Control or Check in the Formation of the Normal Equations. Referring to equations (18), let the sum of the numerical coefficients and the constant term in each equation be represented by s; thus,

$$\begin{aligned}
&a_1+b_1+c_1+\ \ldots\ +l_1=s_1 \\
&a_2+b_2+c_2+\ \ldots\ +l_2=s_2 \\
&\ldots\ldots\ldots\ldots\ldots \\
&a_n+b_n+c_n+\ \ldots\ +l_n=s_n
\end{aligned} \tag{23}$$

To form the first normal equation, as shown in Art. 41, the terms in the left-hand member of each of these equations are multiplied by its weight and its first term or coefficient, namely, w_1a_1, etc., and the resulting products are added, as in (21). Performing this operation at the same time on the right-hand members above, in (23), we have, using the first equation, only, as an illustration:

$$w_1a_1^2+w_1a_1b_1+w_1a_1c_1+\ \ldots\ +w_1a_1l_1=w_1a_1s_1 \tag{24}$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

or, after addition,

$$[wa^2]+[wab]+[wac]+\ \ldots\ +[wal]=[was] \tag{25}$$

Thus, the second member of this equation should equal the sum of the numerical coefficients and the constant term of the first normal equation, which affords a check upon the numerical work of computing these quantities and forming the normal equations. This second member of (25) is therefore called the *sum-term*. For the other normal equations, respectively, it has the form $[wbs]$, $[wcs]$, etc. This check is very important and should always be applied, except, perhaps, in the very shortest problems. Having formed the sum, s, for each of the observation equations, it is treated the same as the other quantities, a, b, c, etc., and when a normal equation is written, its sum-term should equal the algebraic sum of its other numerical quantities. It must be noted, however, that the check may not hold exactly, in the last decimal place, owing to discarded remainders, but this discrepancy will not usually exceed one unit in that place.

44. Symmetry of the Normal Equations. Inspection of the literal forms of the normal equations, in (21) and (22), reveals a symmetry which is useful as an aid to the memory, and which will lessen the labor of computation both in forming the equations and in their solution, as will be shown farther on. This symmetry exists among the coefficients of the unknown quantities with reference to the diagonal line passing downward to the right. On this diagonal will be found those terms which involve the squares of the quantities, a, b, c, etc., as [1] $[waa]$, $[wbb]$, $[wcc]$, etc., or more simply, without weights, $[a^2]$, $[b^2]$, $[c^2]$, etc. These terms being squares, are *always positive*. Then the coefficients in any vertical column occur in the same order as those in the corresponding horizontal row. Thus, in the *third* column and row the order is $[ac]$, $[bc]$, $[cc]$, $[dc]$, etc., c being the *third* of the original coefficients, and the other factors having the original order, a, b, c, d, etc.

45. Formation of the Normal Equations. Aids. The computation of the necessary squares and products for the coefficients in the normal equations is facilitated by the use of special methods as well as tables and mechanical devices. The choice of the method or device will be governed, in general, by the size of the numbers involved and the refinement of the computations.

[1] The squares of a, b, c, etc., are often written as aa, bb, cc, etc., to illustrate this symmetry as well as to avoid the use of exponents.

The *tables* used contain the *logarithms,* the *squares,* or the *products* of numbers. Five-place logarithms are suitable in most work, and four places are often sufficient. Hussey's five-place tables are recommended as very convenient. Barlow's tables of squares, cubes, roots and reciprocals of numbers up to 9,999 are well known and satisfactory. Of the tables of products, Crelle's Rechentafeln, giving the complete products of numbers of three figures each, that is, up to 999 by 999, is probably the most useful, although Zimmermann's and Peters' may more readily be used to obtain products of larger numbers, as they give directly products of numbers of four figures by those of two figures. In computing the coefficients for normal equations by means of tables, the logarithmic method is slowest, the use of squares is better, and the tables of products are usually most satisfactory. In the absence of these last, however, tables of squares may be used in either of two ways for the computation of products, namely, by one of the following formulas:

$$ab = \tfrac{1}{2}[(a+b)^2 - a^2 - b^2] \tag{26}$$

and

$$ab = \tfrac{1}{4}[(a+b)^2 - (a-b)^2] \tag{27}$$

The former requires but one new opening of the tables, as a^2 and b^2 are separately necessary as coefficients.

The *mechanical aids* to computation consist of slide-rules and computing machines. The ordinary 10-in. slide-rule is sufficient for reading products to three significant figures. The Thatcher slide-rule, however, reads directly to four or sometimes five figures and is excellent for solving normal equations as well as forming them. Computing machines are of two types, for addition and for multiplication. We are concerned primarily with the latter, although the former may be used indirectly for multiplication. Of the multiplying machines there are two forms: the Brunsviga type, in which one turn of the crank multiplies by one unit so that to multiply by 43, four turns would be required in one position and three in the next: and the Millionär machine, in which one turn of the crank multiplies by a whole digit, so that but two turns would be required to multiply by 65, one for each digit. If very large numbers are to be multiplied or divided, a computing machine is almost indispensable, but for ordinary work the tables of products and the slide-rule are convenient and sufficient, especially since large numbers are avoided as much as possible.

46. The computation of the coefficients in the normal equations is carried out conveniently in the form of a table in which each quantity involved is shown, with its proper sign, first the given ones and then the computed ones. Then it is highly important that the multiplication of several quantities by the same factor be performed in succession, as this plan in particular is adapted to the use of slide-rules, multiplication tables, and computing machines. Thus, for each observation equation, the factor, wa, would be multiplied into $a, b, c, \ldots l$, and s, in succession, and the products entered in the proper columns of the table, so that the sums of the columns would be the coefficients, $[waa]$, $[wab]$, $[wac]$, . . . $[was]$, of the first normal equation. Next, the factor, wb, would be multiplied into the same quantities, beginning with b, however, as the wab products are included in the preceding series, and the column totals would be coefficients for the second normal equation, and so on. As each normal equation is completed, its coefficients should be tested with the sum-term to assure the computer that the check is satisfied. This would be indicated by a definite check-mark after the sum-term if it checked exactly, or by the cancellation of its last figure with the correct one written above so as to equal the sum of the quantities in the equation.

In the simplest cases, when there are but few observations and two unknown quantities, and when the coefficients are small integers, it may not be worth while to carry out the tabulation for the formation of the normal equations, but it is generally safer to do so, especially when the computer is subject to interruption in his work. It is well, also, to write the algebraic signs for a complete equation before forming and writing the numbers, always writing all positive as well as negative signs.

47. Example of the Direct Formation of Normal Equations. As an illustration of the preceding articles, the normal equations will be formed directly from the following set of observation equations. For simplicity, the weights will be assumed equal.

Observation Equations

$$
\begin{aligned}
+6X+40Y-58.8&=0\\
+4X+32Y-38.3&=0\\
-5X-56Y+43.3&=0\\
-3X-28Y+27.6&=0
\end{aligned}
\qquad (28)
$$

Table for the Formation of the Normal Equations

a	b	l	s	aa	ab	al	as	bb	bl	bs
+6	+40	−58.8	−12.8	+36	+240	−352.8	−76.8	+1600	−2352.0	−512.0
+4	+32	−38.3	− 2.3	+16	+128	−153.2	− 9.2	+1024	−1225.6	− 73.6
−5	−56	+43.3	−17.7	+25	+280	−216.5	+88.5	+3136	−2424.8	+991.2
−3	−28	+27.6	− 3.4	+ 9	+ 84	− 82.8	+10.2	+ 784	− 772.8	+ 95.2
+2	−12	−26.2	−36.2	+86	+732	−805.3	+12.7√	+6544	−6775.2	+500.8√

Normal Equations

$$
\begin{aligned}
+86X+732Y-805.3&=0\quad(+12.7\surd)\\
+732X+6544Y-6775.2&=0\quad(+500.8\surd)
\end{aligned}
\qquad (29)
$$

As there are no discarded remainders, the checks are exactly satisfied.

(The solution of these equations by the methods of algebra gives

$$X = +11.52 \text{ and } Y = -0.25$$

as the best values of X and Y obtainable from the four given observations.)

48. Use of Assumed Approximate Values of the Unknowns. The constant term of the observation equations is sometimes large as compared with the other numerical quantities, and to save labor in the formation of the normal equations, recourse may be had to a scheme similar to that used in Arts. 28 and 33 in the computation of the mean, namely, the use of assumed, approximate values of the unknowns, by which device the constant terms will be reduced considerably in size. For each unknown in the observations, there is substituted its approximate value plus a small correction, as,

$$X = X_0 + x$$

$$Y = Y_0 + y, \quad \text{etc.} \tag{30}$$

where X_0 and Y_0 represent the approximate values and x and y, the small corrections. The approximate values may be obtained by a trial solution of the necessary number of the observation equations, namely, as many as there are unknown quantities.

Thus, in the example of the preceding articles, a solution of the third and fourth of the observation equations results in

$$X = +11.9 \text{ and } Y = -0.29$$

whence we may assume the approximate values,

$$X_0 = +12.0 \text{ and } Y_0 = -0.3$$

Substituting for X and Y, therefore, in equations (28), the quantities

$$X = x + 12.0 \text{ and } Y = y - 0.3$$

we obtain for the first equation,

$$+6(x+12.0)+40(y-0.3)-58.8=0 \tag{31}$$

and for the entire set of observation equations, after simplification,

$$
\begin{aligned}
+6x+40y+1.2&=0\\
+4x+32y+0.1&=0\\
-5x-56y+0.1&=0\\
-3x-28y\qquad&=0
\end{aligned}
\tag{32}
$$

The constant terms have thus been diminished to very small quantities and without the expenditure of much labor, so that the formation of the normal equations will be considerably easier, but in so far only, be it noted, as the terms involving l are concerned. It is obvious that this scheme leaves the *coefficients of the unknowns entirely unaltered,* the only changes being in the constant terms.

49. Adoption of New Unknowns to Equalize Coefficients. When the coefficients of any unknown in the observation equations are consistently large, they may be reduced in size by an artifice similar to that of the preceding article, that is, by substituting for the unknown a new one obtained by multiplying the former by a certain factor.

In the equations (32), for example, the coefficients of y are much larger than those of x and would be easier to handle if they were divided by, say, 20. Therefore, assume

$$y'=20y \quad \text{or} \quad y=\frac{y'}{20} \tag{33}$$

Substituting this value of y in the given equations, and writing the coefficients in columns for simplicity, we have,

$$
\begin{array}{rrrcl}
x & y' & & & \\
+6 & +2.0 & +1.2 & = & 0\\
+4 & +1.6 & +0.1 & & \\
-5 & -2.8 & +0.1 & & \\
-3 & -1.4 & 0 & &
\end{array}
\tag{34}
$$

which are much simpler than the original equations (28), both as regards the formation of the normals and their solution. The normal equations are

$$\begin{array}{cccccc} x & y' & & & & \\ +86 & +36.6 & +7.1 & = & \mathbf{0} & \\ +36.6 & +16.36 & +2.28 & & & \end{array} \tag{35}$$

and their solution results in

$$x=-0.48 \quad \text{and} \quad y'=+0.94 \tag{36}$$

whence,

$$y=\frac{y'}{20}=+0.047$$

Therefore,

$$X=+12.0+x=+12.0-0.48=+11.52$$

and (37)

$$Y=-\ \ 0.3+y=-\ \ 0.3+0.05=-\ \ 0.25$$

The advantages of reducing the size of the coefficients and constants before forming and solving the normal equations, is less in such a short problem as the one just solved than in the ones which contain more unknown quantities and larger series of observations. It is generally advisable, however, even in the shorter problems, to diminish the coefficients and constants to a size which will be convenient for computation, and to equalize them to some extent, at least, by using whole numbers for the approximate values and the factors.

50. Example: Time by Star Transits. Let us consider, as an example of indirect observations, the determination of time by observed transits of stars on the meridian, using an astronomical transit instrument. The times when each star is seen to cross the successive threads are recorded by the observer, himself, as he carries the beats of the chronometer in his mind. The mean of these times is taken as the time when the star crossed the line of sight of the instrument. It is then corrected for diurnal aberration, the rate of the chronometer, and the inclination of the horizontal axis of the instrument as determined from the readings of the striding level. The resulting time, θ', is subtracted from the right ascension, α, of the star (which is the correct sidereal time when the star crosses the *meridian*) and the difference,

$\alpha-\theta'$, according to the usual notation, is therefore made up of the chronometer correction, $\Delta\theta$, which is the quantity really desired from the observations, and the corrections for azimuth, Aa, and collimation, Cc, according to the formula,

$$Aa+Cc+\Delta\theta-(\alpha-\theta')=0 \tag{38}$$

in which A and C are the known azimuth and collimation factors, and a, c, and $\Delta\theta$ are respectively the azimuth and collimation constants and the chronometer correction, which are the three unknowns of the problem and which, therefore, will be represented by x, y, and z. Each star thus furnishes an observation equation of the form,

$$Ax+Cy+z-(\alpha-\theta')=0 \qquad \text{Wt. } w \tag{39}$$

the weight being determined from the star's declination, as stated in Art. 32. The given data for the nine stars observed are:

A	C	$\alpha-\theta'$	w
+0.15	+1.22	−9m. 06.81s.	0.9
+0.72	+1.00	06.98	1.0
−1.66	+3.30	07.69	0.2
−0.17	+1.52	06.49	0.7
+0.09	−1.28	06.08	0.8
−0.18	−1.53	06.32	0.7
+0.80	−1.02	06.27	1.0
+4.25	+4.92	07.47	0.1
+0.60	−1.01	06.00	1.0

(40)

As the values of $\alpha-\theta'$ are nearly equal, and the coefficients of z are unity, it is evident that the value of z will approximate to $\alpha-\theta'$. Therefore, let

$$z=z'-9\text{m}.06.00\text{s}. \tag{41}$$

and the form of the typical equation becomes

$$Ax+Cy+z'+l=0 \qquad \text{Wt. } w \tag{42}$$

in which $l=-9\text{m}.06.00\text{s}.-(\alpha-\theta')$. The modified observation equations, therefore, are

$$+0.15x+1.22y+z'+0.81=0 \quad (\text{Wt. } 0.9), \quad \text{etc.,}$$

or, arranged in tabular form,

x	y	z'	(l)		(Wt.)
+0.15	+1.22	+1	+0.81	= 0	0.9
+0.72	+1.00	+1	+0.98		1.0
−1.66	+3.30	+1	+1.69		0.2
−0.17	+1.52	+1	+0.49		0.7
+0.09	−1.28	+1	+0.08		0.8
−0.18	−1.53	+1	+0.32		0.7
+0.80	−1.02	+1	+0.27		1.0
+4.25	+4.92	+1	+1.47		0.1
+0.60	−1.01	+1	0		1.0

(43)

The quantities in the successive columns are the a, b, c, l, and w, of the observation equations, so that it is unnecessary to tabulate them again.

Forming the normal equations from the typical ones in (21) by means of a tabulation similar to that of Art. 47, we have,

x	y	z'	(l)		(s)
+3.94	+ 0.38	+2.18	+1.00	= 0	+ 7.50
+0.38	+13.56	+0.19	+3.53	= 0	+17.66
+2.18	+ 0.19	+6.40	+3.09	= 0	+11.86

(44)

the solution of which, as algebraic simultaneous equations, gives

$$x = +0.043\text{s.},\qquad y = -0.261\text{s.},\qquad z' = -0.491\text{s.},$$

whence, from (41),

$$z = \Delta\theta = z' - 9\text{m}.06.00\text{s.} = -9\text{m}.06.49\text{s.},$$

so that $a = +0.043$s., $c = -0.261$s., and $\Delta\theta = -9$m.06.49s., represent the best or most probable values of the three unknowns which can be obtained from the given observations.

51. General Application of the Method. The process outlined in this chapter for the adjustment of indirect observations may be applied to any set of linear simultaneous equations whose number is greater than that of the unknown quantities, although they may not be observation equations. Sometimes such a series of equations may result from computations or from theoretical assumptions. Inasmuch, however, as the present method of adjustment depends upon the assumptions as to the occurrence of error, stated in Art. 18, the use of the method for the adjustment of quantities other than those resulting from observations may be justified only by the absence of a better scheme.

However, any other method is likely to be more laborious than this one, if it takes into account all the given data. For example, suppose the simplest case of three given equations involving two unknowns. If ignorant of the adjustment by means of Least Squares, but desirous of utilizing all of the given equations because there is no way of telling which one could be discarded with least effect, the computer might reasonably select all possible combinations of the three equations, two at a time, namely, three, and solve each of the three pairs independently by algebraic methods, thus obtaining three different values for each of the unknowns, of which he would probably take the mean as the best value within his knowledge. Certainly, the formation and solution of two normal equations would be much easier than such a process.

CHAPTER IV

SOLUTION OF NORMAL EQUATIONS

52. Methods of Elimination. As simple, simultaneous equations of the first degree, the normal equations may be solved by any of the ordinary algebraic methods of elimination; by addition or subtraction, by substitution, or by comparison. In fact, these methods are satisfactory when there are but two equations to be solved. But in larger sets, of three or more, it is possible to shorten the numerical work by taking advantage of the peculiar symmetry which all normal equations possess, as was pointed out in Art. 44. It is much easier to solve a set of normal equations than a set of ordinary, simultaneous equations of the same number which do not have this symmetry.

53. The Gauss Method of Substitution has been for a century the basis of the special methods for the solution of normal equations. Its notation is convenient, and in its general, literal form it is given in nearly every work on Least Squares. However, it has been modified and improved in various ways, particularly by Mr. M. H. Doolittle, formerly a computer in the U. S. Coast and Geodetic Survey, and, in the effort to confine ourselves to a single method which shall be the most generally useful one for our purposes, we shall omit the detailed explanation of the Gauss process.

54. Requirements of a Good Method. It is important that the method to be adopted be as universally useful as possible, in both short and long problems, although modifications may be convenient to adapt it to special or peculiar cases. The various steps in the elimination should be identical, so that the work may be performed mechanically, to a great extent, thereby avoiding mistakes. The method should be as short as possible so as to avoid unnecessary work. And finally, checks should be available at frequent intervals throughout the computation in order that errors may be discovered and corrected without a great deal of recomputation. All of these qualities should be borne in mind and utilized as far as possible in every solution. It is believed that the Abridged Method explained below fulfills these requirements and that it will be readily understood.[1]

55. Algebraic Elimination by Addition. Let us undertake the solution of a simple set of normal equations in order that the steps we shall take in the process may be clearly understood. The method of elimination by addition will be used, although arranged in a certain form to illustrate the shorter method which is to follow. The given normal equations, with coefficients arranged in columns, are:

	x	y	z	(l)	
(I)	$+6$	-2	$+3$	$+2$	$=0$
(2)	-2	$+3$	-4	-3	$=0$
(3)	$+3$	-4	$+3$	$+1$	$=0$

(45)

For purposes of explanation, the equations are numbered at the left, but for a reason which will appear later, the first is given the Roman numeral (I). First, we eliminate x between the first and second equations, by multiplying the first by such a quantity or factor, as will make its first term equal to that of the second equation with the opposite sign, and then adding the two. This factor will be the quotient of the first term of the second equation with its sign changed, by the first term of the first equation, that is, $+2/6$. Thus, indicating on the right the steps taken, we write down equation (2) and under it the first equation multiplied by $+2/6$:

	x	y	z	(l)		
(2)	-2	$+3$	-4	-3	$=0$	
(4)	$+2.0$	-0.7	$+1.0$	$+0.7$	$=0$	(I)×($+2/6$)
(II)	0	$+2.3$	-3.0	-2.3	$=0$	(2)+(4)

(46)

[1] Published by M. H. Doolittle in C. & G. Survey Report, 1878, App. 8.

This equation, resulting from the elimination of the first unknown is called a *First Derived Equation*, and is given the Roman numeral (II) as marking the completion of a whole step in the process.

Next, x is eliminated in the same way between the first and third equations, by multiplying the first by such a factor as will make its first term equal to that of the third equation with its sign changed, and adding the two equations. As before, the factor will be the quotient of the latter first term with the reversed sign by the former one, or $-3/6$. Writing (3) first,

	x	y	z	(l)		
(3)	+3	−4	+3	+1	=0	
(5)	−3.0	+1.0	−1.5	−1.0	=0	(I)×(−3/6)
(6)	0	−3.0	+1.5	0	=0	(3)+(5)

(47)

Equation (6) is like (II) in having no x-term, that is, it may also be called a First Derived Equation. Therefore, y can be eliminated from these two equations in the same manner as x was eliminated above. Multiplying (II) by the first term of (6) with its sign changed, divided by the first term of (II), and adding the resulting equation to (6), we have, as before, writing the second of the two equations first:

	y	z	(l)		
(6)	−3.0	+1.5	0	=0	
(7)	+3.0	−3.9	−3.0	=0	(II)×(+3.0/2.3)
(III)	0	−2.4	−3.0	=0	(6)+(7)

(48)

This completes the second step in the process, the second unknown has been eliminated and this last equation is therefore called a *Second Derived Equation* and given the next Roman numeral (III). The elimination is now complete and the last equation may be written

$$-2.4z-3.0=0$$

giving the value of z directly, as $z = +3.0/-2.4 = -1.25$. Substituting this value in (II) gives

$$y = \frac{+3.0z + 2.3}{+2.3} = \frac{-3.75 + 2.3}{+2.3} = \frac{-1.45}{+2.3} = -0.63$$

Then, from (I), (49)

$$x = \frac{+2y - 3z - 2}{+6} = \frac{-1.26 + 3.75 - 2}{+6} = \frac{+0.49}{+6} = +0.08$$

Some properties of this method of solution will now be considered.

56. Symmetry among the Derived Equations. The First Derived Equations resulting from the elimination of the first unknown, x, between the first normal equation and each of the others in succession, will be one less, in number, than the unknowns; therefore, in the above example they are two, namely,

	y	z	(l)	
(II)	+2.3	−3.0	−2.3	=0
(6)	−3.0	+1.5	0	=0

(50)

They evidently are symmetrical about the diagonal in the same respect as the normal equations, that is, the second coefficient in the first row (−3.0) is the same as the second coefficient in the first column (−3.0). This symmetry exists likewise in all sets of first derived equations, whatever their number, as may be proved by carrying out a complete solution of the typical normal equations with their literal coefficients. In this example, however, the reason for the equality of these two coefficients may be seen by indicating the operations through which they were obtained, using certain symmetrical terms in the normal equations. Thus,

Second Coeff. of (II) $= -4 - (-2/6 \times +3) = -3.0$

First Coeff. of (6) $\quad = -4 - (+3/6 \times -2) = -3.0$ (51)

The two -4's are symmetrical in the normal equations, as also are the two -2's and the two $+3$'s, while the $+6$ is the same in both cases.

Likewise, the Second Derived Equations, resulting from the elimination of y from the first derived equations, are symmetrical among themselves, and so on with successive sets of derived equations in the solution of a large number of normal equations by this method of elimination.

57. Omission of Redundant Terms. (*a*) In each first derived equation, x has been eliminated; that is, it has the coefficient zero. Therefore, it is unnecessary to write the coefficients in the x-column at all during the elimination of x, as in (46) and (47), as we know that they will add up to zero if the work is correct, and anyway, there will be other and better checks on the correctness of the work. Similarly, the y-column may be omitted during the elimination of y, as in (48), and so on. However, the sum of the remaining terms in each equation will not now equal zero, except in the derived equations, where the omitted coefficients are always zero. This will deprive us of the equation-checks except in the derived equations, but these will still be sufficiently close together to control the computation.

(*b*) By transposing all the terms of each equation into one member, as was done in the above example, we may omit the symbols, " $=0$," from each equation. As just stated, however, these must be understood, in the cases of the derived equations, as if written.

(*c*) Even the original normal equations may be simplified by the omission of all the terms lying below the diagonal, these being symmetrical to the ones above the diagonal. Thus, in the normal equations, (45),

	x	y	z	(l)	
(1)	$+6$	-2	$+3$	$+2$	$\neq 0$
(2)	-2	$+3$	-4	-3	$\neq 0$
(3)	$+3$	-4	$+3$	$+1$	$\neq 0$

(52)

the canceled terms would be omitted. To read the original equations, then, the omitted portion of each *row* must be replaced by the symmetrical quantities in the corresponding *column*. For example, the second equation is begun in the second column and read downward to the diagonal and then horizontally to the right along its own row, retaining, however, the original order of the unknowns, as $-2x+3y-4z-3=0$. Similarly, the third equation is begun in the third column, read downward to the diagonal and continued along the third row as usual: $+3x-4y+3z+1.0=0$. The simplified form of the equations, then, would be:

	x	y	z	l
(1)	+6	−2	+3	+2
(2)		+3	−4	−3
(3)			+3	+1

(53)

It is evident that the original *order* of the equations must never be changed if these abbreviations are to be used, as the symmetry would then be destroyed.

(*d*) In the second step of the elimination, it is unnecessary to write equation (6) either in (47) or (48), but equation (7) may be written directly below (5) in (47), the x- and y-columns being omitted as above, and the three lines (3), (5), and (7) added, to form (III). The factor for obtaining (7) from (II), namely, $-(-3.0/2.3)$, is the second term of (II) with its sign changed, divided by the first term of (II), owing to the symmetry of the derived equations as shown in Art. 56. Thus, (47) and (48) may be combined into:

	z	l	
(3)	+3	+1	
(5)	−1.5	−1.0	(I)×(−3/6)
(7)	−3.9	−3.0	(II)×(+3.0/2.3)
(III)	−2.4	−3.0	(3)+(5)+(7)

(54)

58. The Series of Derived Equations. Upon inspection of (46) and (54), it will be seen that (II) is derived from (2) and (I), and that (III) is derived from (3), (I), and (II). If there were a fourth unknown and four normals, the derived equation (IV) would be derived from the fourth normal equation and (I), (II), and (III), and so on. Here, then, lies the reason for giving to the first normal equation the Roman numeral (I); it is associated with the derived equations in each step of the elimination. Therefore, in writing a list of the derived equations, this equation is written first, and is referred to as one of them. Such a list, in order, has the property that each equation is complete and begins with the second unknown of the preceding one, so that the series is used for determining the successive unknowns in their reverse order when the elimination has been completed, by substitution back through them.

59. Control or Check in the Solution of the Normal Equations. The check on the formation of the normal equations, explained in Arts. 43 and 47, may be continued through the process of elimination so as to test the correctness of the computation at frequent intervals. If the sum-term of each of the normal equations be subjected to the same operations as its other terms, the resulting modified sum-term will be equal to the sum of the corresponding series of other terms. Moreover, this relation will persist when several equations have been added or subtracted, the sum of all the sum-terms being equal to the sum of all the other terms. Thus, the sum-terms which were used to check the formation of the normal equations may be used during their solution to test the correctness of an equation at any stage of the work. As was pointed out in (*a*) of the last article, however, the omission of redundant terms leaves the derived equations as the only complete ones in the elimination. Therefore, this sum-check being applicable to each derived equation as it is formed, should be taken advantage of in every case.

Since the check applies only to complete equations, the coefficients of the normal equations must be read down and to the right as shown in the last article, when the simplified form is used.

In the above example, then, the complete statement of the normal equations in the simpler form, with their check-terms, is:

	x	y	z	(l)	(s)
(1)	+6	−2	+3	+2	+9
(2)		+3	−4	−3	−6
(3)			+3	+1	+3

(55)

60. Elimination by the Abridged Method. This set of equations will now be solved in accordance with the devices explained in the preceding articles for abridging the various operations as much as possible. A comparison of this solution with the direct, algebraic one given in Art. 55, will illustrate the different steps and the saving of labor.

	x	y	z	(l)	(s)	
(I)	+6	−2	+3	+2	+9	
(2)		+3	−4	−3	−6	Normal Equations
(3)			+3	+1	+3	
(2)		+3	−4	−3	−6	
(4)		−0.7	+1.0	+0.7	+3.0	(I)×(+2/6)
(II)		+2.3	−3.0	−2.3	−3.0✓	(2)+(4)
(3)			+3	+1	+3	
(5)			−1.5	−1.0	−4.5	(I)×(−3/6)
(6)			−3.9	−3.0	−3.9	(II)×(+3.0/2.3)
(III)			−2.4	−3.0	−5.1✓	(3)+(5)+(6)

The process may be outlined as follows: Write (2); follow the left-hand column of (2) up to (I) and find −2; change its sign and divide by the first term of (I), giving +2/6; multiply this factor into the terms of (I), beginning with the same left-hand column of (2), and writing the results in line (4) under the corre sponding ones in (2); add (2) and (4) to obtain (II). Next, write (3); follow its left-hand column up to (I) and find $+3$; change its sign and divide by the first term of (I), giving $-3/6$; multiply this factor into the terms of (I), beginning with the left-hand column of (3), writing the products in their proper columns in line (5), under (3). Again, follow the same left-hand column of (3) up to (II) and find -3.0; change its sign and divide by the first term of (II), giving $+3.0/2.3$; multiply this factor into the terms of (II), beginning with the left-hand column of (3), writing the results in line (6), below (5), and in their proper columns; add (3), (5), and (6), to obtain (III), as the second and last step in the elimination. If there were four normal equations, the fourth would now be written, beginning with the fourth column; under it would be written the products of the terms of (I), beginning with the fourth, by a factor consisting of the fourth term of (I) with its sign changed, divided by its first term; under this line would be written the products of the terms of (II), beginning with the third, by a factor consisting of this third term with changed sign, divided by the first term of (II); and finally, under this line would be written the products of the terms of (III), beginning with the second, by a factor consisting of this second term with its sign changed, divided by the first term of (III); whence the sum of the *four* lines thus obtained would be (IV). This procedure could be continued through any number of equations.

61. The mechanical character of this scheme of elimination is apparent from the foregoing explanation. Each of the main steps accomplishes the elimination of one unknown more than the preceding step did, and results in the next derived equation. Each step consists of the sum of its normal equation and as many others as there are derived equations already formed, including (I); so that the successive steps embrace the sums of two, three, four, five, or more lines, up to the number of unknowns involved in the problem. For each step, the numerators of the factors, with opposite signs, are found in one column, namely, the one containing the first term of the normal equation as written, that is, the one corresponding to that equation, as third column for third equation, etc.; the denominators are the first terms of the corresponding derived equations.

62. Notes and Suggestions. The arrangement of the work in columns is essential to mechanical efficiency and "Data Sheets" are convenient for this purpose. Ruled horizontal lines including each derived equation make it prominent for quick reference. By writing the algebraic signs of each line of products before writing the numbers themselves, errors in sign may be avoided to a great extent. In each line, all the signs will be the same as those of the corresponding derived equation, or all opposite to them. It will be noted that the *first term* in each of these lines of products is *always negative* owing to changing the sign in the numerator of each factor; this, also, affords a check on the signs. Unavoidable discrepancies in the last figure of the check-term, due to remainders, should be removed by arbitrarily correcting the check-term before proceeding with the next step in the elimination; this is best done by drawing a line through the erroneous figure and writing the correct one just above it. If the check is exactly satisfied, it should be as carefully noted with a check-mark in order to avoid uncertainty.

63. Values of the Unknowns. The process of elimination having been completed and the derived equations checked as formed, it remains to determine the last unknown from the last equation and to substitute back in the preceding derived equations in reverse order, to obtain the other unknowns; x being finally determined from (I). If there be many equations, this process may be facilitated by tabulating the products instead of indicating the work as in (49). A table is begun for each unknown by writing first, with changed sign, the constant term of the derived equation from which that unknown is to be obtained. Below this are placed in succession the products of the unknowns, as computed, by their respective coefficients, with signs changed, in that equation. The sum of these quantities divided by the first coefficient in the equation gives the value of that unknown. The advantage lies in the fact that each unknown, as computed, is multiplied into all of its coefficients in the preceding derived equations, in succession, rather than separately as needed for each case; thus, a slide-rule, a multiplication table, or a machine can be used with profit. Applying this arrangement to the problem in Art. 60, we have:

	x	y	z
Constants	-2.0	$+2.3$	$\dfrac{+3.0}{-2.4} =$
z-terms	$+3.75$	-3.75	
y-terms	-1.26		$-1.25 = z$
		$\dfrac{-1.45}{+2.3} =$	
	$\dfrac{+0.49}{+6.0} =$		
		$-0.63 = y$	
	$+0.08 = x$		

It is often convenient to write these little tables in the proper columns just to the left of the computation of the derived equations, that is, left of the elimination, as it is not necessary to write the numbers of the various equations after the method has been learned.

64. Final Check of the Unknowns. The use of the check-terms ends with the formation of the last derived equation. The entire solution, however, including the values of the unknowns, may be checked by substituting those values in the normal equations other than the first. If there were no neglected remainders in the computation, all these equations should be exactly satisfied. Therefore, in any actual case, the discrepancies should be very small.

When several equations are to be tested at once, the method of tabulation explained in the last article may be used to advantage. Obviously, the sum of each table should be very nearly zero if no mistake has been made. This will be illustrated in a later chapter.

65. Refinement of the Computations. The number of decimal places to be retained throughout the elimination will depend upon the number desired in the resulting unknowns. The labor of solution increases rapidly with the size of the quantities involved so that it is very important to keep them as small as practicable. The discrepancies due to neglected remainders during the elimination will seldom amount to more than one or two in the last place and these will be revealed by the checks. Similar ones will occur in the values of the unknowns, resulting in their failure to check exactly when substituted in the original normal equations. However, as the final values of the unknowns must be regarded as but approximations to the correct values, which, of course, are unattainable, it cannot be objectionable to alter the last figure of an unknown arbitrarily, to make it check or to make it consistent with the others, and this is sometimes necessary. Therefore, it is unwise to carry the whole computation one or two places farther merely to secure an exact check in a certain place without forcing it.

As a general rule, it is well to carry the observations two places beyond the last one which is regarded as known with certainty. For example, each reading will have its last figure the result of estimation, to some extent, the preceding one being certain; then the mean of several readings would be carried one place farther. This should determine the degree of refinement to which the normal equations and the elimination should be carried, the coefficients of the observation equations being modified as shown in Arts. 48 and 49 so as to be consistent in size. The unknowns may then be carried out one place farther, the last figure to be retained or rounded-off to the preceding one as preferred. However, this is largely a matter of judgment derived from experience. The beginner is too apt to carry his work farther than is justified by the precision of the observations. He may be guided by the rule to carry the computations one place farther than the given data; this is ample.

66. Mechanical Aids in the Solution. We have seen in Art. 45 how the formation of the normal equations may be facilitated by the use of tables and mechanical devices. In the solution of the equations, these articles are even more useful, perhaps, especially the slide-rules, as they admit of multiplying a series of numbers by the quotient of two other numbers at one setting of the rule. The Thatcher, in particular, is very convenient, and is good for four significant figures, but the ordinary 10-in. rule is sufficient when but three figures are used and is commonly at hand. The computing machines, while necessary for large numbers are less advantageous for small ones, but they have the great advantage over the slide-rules of causing little or no straining of the eyes.

CHAPTER V

OBSERVATIONS OF DEPENDENT QUANTITIES: CONDITIONED OBSERVATIONS

67. Dependent Quantities. In the preceding chapters, the quantities observed or determined from the observations have been independent, that is, any one or more might vary without causing a corresponding change in the others. Thus, in the determination of time by star transits in Art. 50, the constants of the transit instrument cannot be affected by any change in the chronometer correction. Now, however, we shall consider a different class of quantities, and one which is of particular importance to engineers, inasmuch as it includes their most complex, but at the same time, most useful, problems in the adjustment of observations. In this second division of the subject, the observed quantities are not independent of one another, but are inter-related by certain theoretical requirements, called *Conditions*, which their adjusted or adopted values must *rigidly satisfy*. The adjustment, then, consists in determining the best set of values for the observed quantities which shall exactly satisfy the prescribed conditions.

For example, if the three angles of a plane triangle be measured with a protractor, they must be so adjusted that the sum will be exactly 180°. Or, if the three angles of a triangle in the field be measured with a transit or theodolite, they must be adjusted by the application of a small correction to each, so that the sum of the adjusted values will be 180° plus the spherical excess.[1] Also,

[1] The earth is approximately spheroidal but the figures in triangulation are considered as spherical for convenience in computation. The observed horizontal angles, then, are those of *spherical* triangles, since the plumb-lines at the different stations are convergent and the horizontal planes of the angles are neither coincident nor parallel. Very small triangles, however, may be considered plane, as the spherical excess is but 1″ in a triangle containing 75 square miles.

the horizontal angles completing the horizon at a station must be adjusted so that the sum will be 360°; and the differences of elevation in a closed circuit of levels must be adjusted so that their algebraic sum will be zero, when proceeding continuously around the figure, that is, clockwise or counterclockwise.

68. The observations to be adjusted will have been made independently, as a rule, as in the case of a circuit of levels made up of several differences of elevation between successive bench-marks, each difference of elevation being determined independently of the others. However, so-called "observations," entering into an adjustment, may never have been actually observed but may be the results of computation or of a previous adjustment of actual observations. For example, an angle of a triangle may have been determined by the addition or subtraction of two or more observed angles, or from a local adjustment of the angles at that station.[1] Also, as stated in Art. 26, each observation may be the result of several elemental observations or readings; in fact, this is usually the case with dependent quantities. Generally, too, these observations are direct ones. In any event, however, they will be adjusted as direct observations of dependent quantities, as this is the most convenient and practical method.

69. The weights, in the general case, will be unequal, of course. They are obtained as indicated in Art. 32, from the number of observations, from theory, or by estimation. They may be determined from the nature of the observed quantities, independently of the observations themselves, although subject to modification in every case when the circumstances are unusual. The basis of weights in observations of angles is usually the number of observations; in leveling, the lengths of the lines, the number of instrument stations, etc., may indicate the weights.

70. Conditions. The nature of the conditions which are to be satisfied by the adjusted values of the observed quantities will depend upon the character of the problem. The only limitations

[1] It will be seen later on that it is often convenient to make two or more small, partial adjustments instead of a single large one, so that it frequently happens that the given data to be adjusted are the results of a previous adjustment.

upon them are: (1) that their number must be *less* than that of the observations, as otherwise a sufficient number of them could be solved as simultaneous equations so as to determine the unknowns directly, without an adjustment; and (2) that they must be independent of one another, that is, no condition may be included twice in the same series. Furthermore, the correctness of the conditions is not essential to the adjustment, itself, as this can be carried out so as to force the unknowns to satisfy almost any arbitrary or unreasonable condition; but a *correct* adjustment requires that the conditions be correctly stated. If an error be made in the statement of a condition, and the proper method of adjustment be used, the unknowns would satisfy the erroneous condition, and the error might not be discovered until, as a final check, the adjusted values were tested by substitution in the original conditions. Therefore, it is well to exercise great care in the statement of each condition, and to be sure that all of the necessary conditions, but no others, be included in an adjustment.

71. Number of Conditions. It is evident, in general, that a certain number of observations would be necessary for the determination of a certain number of quantities, if the observations were strictly correct,—ideal. If extra observations are made, beyond this necessary number, each of these would furnish a check upon the work, that is, a condition to be satisfied. The *rule*, then, could be stated that the number of independent conditions of a certain kind, involved in a given series of observations, would be equal to the number of *extra observations* of the corresponding kind, that is, the excess over the necessary number of ideal, correct observations.

Let Fig. 2 represent a system of levels connecting the bench-marks, A, B, C, D, E, and F, the numbers in parentheses representing the lines over which the differences of elevation are observed. If the observations were absolutely correct, the differences of elevation would be completely determined by the lines, (1), (2), (3), (4), and (5). Then if (6) were added, it would furnish one check, and the condition that the whole outer circuit should close to zero, if the signs of the separate lines were so changed, if necessary, as to indicate running continuously around the figure. By adding the line (7), between C and F, another check is obtained, with the corresponding condition that the circuit A–B–C–F should close, or that the remainder of the figure, namely, C–D–E–F, should close. Having taken the closure of the whole figure as the first condition, only one of the two smaller

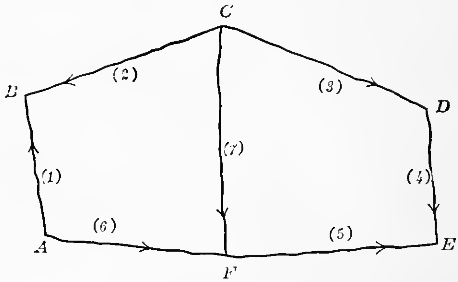


Fig. 2. System of Levels

circuits may be used as an independent condition, since the other small circuit would then necessarily close, being the difference of the other two circuits. Thus, as stated in the above rule, each extra observation gives one independent condition. It is obvious that *any five* lines connecting the *six* benchmarks could be considered as the original, necessary observations, and that *any two* of the three circuits could be used for the two conditions.

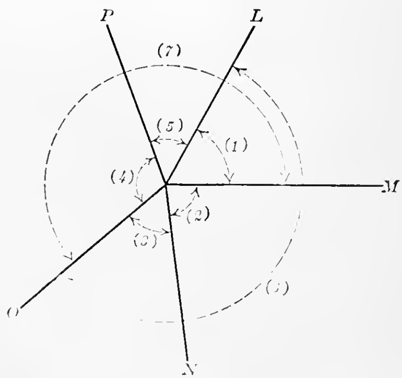


Fig. 3. Horizontal Angles at a Station

As another example, let Fig. 3 represent a series of horizontal angles around a station, connecting the five signals, L–M–N–O–P, each angle observed being indicated by a number in parentheses and a corresponding arc. Four angles would be sufficient to connect the five signals, so that there are three extra observations and, therefore, three independent conditions. If (1), (2), (5), and (6), be regarded as the necessary angles, (3) would give the condition that (1)+(2)+(3)−(6) should equal zero; (4) would complete the horizon with the requirement that (4)+(5)+(6) should equal 360°; and (7) would close the horizon, likewise, with (6)−(1). Different combinations could be used, as well, for the three conditions, such as (1)+(4)+(5)−(7)=0, etc., and these would be independent if each of the three extra observations were used in *one, and only one, condition.* It is easily seen that the Number of conditions=

(Number of angles observed)−(Number of signals)+1.

72. Statement of Conditions. Although the conditions, as functions of the observed quantities, might be very complex in form, involving the higher powers, etc., still in the problems with which the engineer is usually concerned, they are of the linear form or easily reducible to that form. Therefore, we shall confine our attention to these simpler conditions and consider them all as in the linear form.

The conditions express the relations which must be rigidly satisfied by the final, adjusted values of the observed quantities. Let these best, adopted values be represented by $V_1, V_2, V_3, \ldots V_n$; the corresponding observations, by $M_1, M_2, M_3, \ldots M_n$; and the small corrections to be added to the observations to obtain the adjusted values by, $v_1, v_2, v_3, \ldots v_n$, in which n is the number of observations, which is also, in this case, the number of observed quantities. Then $V_1=M_1+v_1$, $V_2=M_2+v_2$, etc. The original conditions will be stated in the following *Condition Equations:*

$$
\begin{aligned}
&a_1V_1+a_2V_2+a_3V_3+\ \ldots\ +a_nV_n+a_0=0\\
&b_1V_1+b_2V_2+b_3V_3+\ \ldots\ +b_nV_n+b_0=0\qquad(56)\\
&c_1V_1+c_2V_2+c_3V_3+\ \ldots\ +c_nV_n+c_0=0\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots
\end{aligned}
$$

in which the a's, b's, c's, etc., are known constants. There will be as many of these equations, of course, as there are independent conditions in the problem, and as many V's as observed quantities.

As the observed values approximate closely to the V's in all observations which are carefully made, they will *nearly* satisfy the conditions (56). Therefore, substituting M's for V's in (56) will result in a small quantity, q, instead of zero, as the value of each condition function, thus:

$$\begin{aligned} a_1M_1+a_2M_2+a_3M_3+\ldots+a_nM_n+a_0&=q_1\\ b_1M_1+b_2M_2+b_3M_3+\ldots+b_nM_n+b_0&=q_2\\ c_1M_1+c_2M_2+c_3M_3+\ldots+c_nM_n+c_0&=q_3\\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots& \end{aligned} \tag{57}$$

q being the amount by which the observations fail to satisfy a condition equation, that is, it is the *closure error* of each condition equation.

Now substitute for V, in the equations (56), the value $M+v$, and we have,

$$\begin{aligned} a_1v_1+a_2v_2+\ldots+a_nv_n+(a_1M_1+a_2M_2+\ldots+a_nM_n+a_0)&=0\\ b_1v_1+b_2v_2+\ldots+b_nv_n+(b_1M_1+b_2M_2+\ldots+b_nM_n+b_0)&=0\\ c_1v_1+c_2v_2+\ldots+c_nv_n+(c_1M_1+c_2M_2+\ldots+c_nM_n+c_0)&=0\\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots& \end{aligned} \tag{58}$$

in which the parenthetical expressions are the values of q in (57). Therefore, the equations (58) take the form,

$$\begin{aligned} a_1v_1+a_2v_2+\ldots+a_nv_n+q_1&=0\\ b_1v_1+b_2v_2+\ldots+b_nv_n+q_2&=0\\ c_1v_1+c_2v_2+\ldots+c_nv_n+q_3&=0\\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots& \end{aligned} \tag{59}$$

These are the *Reduced Condition Equations*. They state the required relation between the *corrections* to the observations and the closure errors of the original conditions. These corrections are the *unknowns* which are to be obtained as a result of the adjustment. The reduced conditions thus involve much smaller quantities than the original ones, (56), and are more convenient to handle.

Comparing the two sets of equations, (56) and (59), we note that they are alike in form but differ only in the substitution of the small v's for the V's and the q's for the constant terms, a_0, b_0, c_0, etc. It is usually convenient, therefore, to write the conditions in the reduced form in the first place, especially as the constants, a_0, b_0, c_0, etc., are zero in most of our problems. However, if the original equations be omitted, the sign of q should be determined with great care. It should be the same as the error of closure of that condition and opposite to that of the corrections, in general. For example, if the sum of the angles closing the horizon at a station be greater than 360°, q would be positive, since the corrections to the angles, generally, would be negative so as to reduce their sum to 360°. For the beginner, nevertheless, it is safer to write the original conditions first, so as to avoid this difficulty with the signs. It should be noted that if an adjustment were carried out completely with the signs of *all* the q's incorrect, it would result in a set of corrections having the wrong signs throughout, which could be changed without altering the adjustment computation in the least.

73. Adjustment by the Method of Correlates. The final, adjusted values of the observed quantities must exactly satisfy the prescribed conditions of the problem, and must be, moreover, the best, or most probable, values, according to the Theory of Least Squares, which will so satisfy them. Therefore, the sum of the weighted squares of the corrections, which have the nature of residuals, must be a minimum, as in Art. 34. That is,

$$[wv^2] = w_1v_1^2 + w_2v_2^2 + w_3v_3^2 + \ldots + w_nv_n^2 = \text{a minimum} \qquad (9)$$

which must be satisfied simultaneously with the conditions (59).

Multiplying the condition equations of (59) in succession by the factors, $-2A$, $-2B$, $-2C$, etc., respectively,

$$\begin{aligned}
&-2a_1Av_1 - 2a_2Av_2 - 2a_3Av_3 - \ldots - 2a_nAv_n - 2q_1A = 0 \\
&-2b_1Bv_1 - 2b_2Bv_2 - 2b_3Bv_3 - \ldots - 2b_nBv_n - 2q_2B = 0 \qquad (60) \\
&-2c_1Cv_1 - 2c_2Cv_2 - 2c_3Cv_3 - \ldots - 2c_nCv_n - 2q_3C = 0 \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots
\end{aligned}$$

Adding these equations to (9) and collecting the coefficients of the separate v's, we have the requirement that

$$\begin{aligned}
&w_1v_1^2-2v_1(a_1A+b_1B+c_1C+\dots)+\\
&+w_2v_2^2-2v_2(a_2A+b_2B+c_2C+\dots)+\\
&+w_3v_3^2-2v_3(a_3A+b_3B+c_3C+\dots)+\\
&+\dots\dots\dots\dots\dots\dots+\\
&+w_nv_n^2-2v_n(a_nA+b_nB+c_nC+\dots)\\
&\qquad -2\quad(q_1A+q_2B+q_3C+\dots)=\text{a minimum}
\end{aligned}\tag{61}$$

For the minimum, the derivative of this expression with respect to each of the v's must be placed equal to zero. Therefore,

$$\begin{aligned}
&2w_1v_1-2(a_1A+b_1B+c_1C+\dots)=0\\
&2w_2v_2-2(a_2A+b_2B+c_2C+\dots)=0\\
&\dots\dots\dots\dots\dots\dots\\
&2w_nv_n-2(a_nA+b_nB+c_nC+\dots)=0
\end{aligned}\tag{62}$$

whence,

$$\begin{aligned}
&v_1=\frac{1}{w_1}(a_1A+b_1B+c_1C+\dots)\\
&v_2=\frac{1}{w_2}(a_2A+b_2B+c_2C+\dots)\\
&\dots\dots\dots\dots\dots\dots\\
&v_n=\frac{1}{w_n}(a_nA+b_nB+c_nC+\dots)
\end{aligned}\tag{63}$$

Substituting these values of the v's in the condition equations (59), and combining the coefficients of A, B, C, etc., we obtain the *Normal Equations:*

$$\begin{aligned}
&\left[\frac{aa}{w}\right]A+\left[\frac{ab}{w}\right]B+\left[\frac{ac}{w}\right]C+\dots+q_1=0\\
&\left[\frac{ab}{w}\right]A+\left[\frac{bb}{w}\right]B+\left[\frac{bc}{w}\right]C+\dots+q_2=0\\
&\left[\frac{ac}{w}\right]A+\left[\frac{bc}{w}\right]B+\left[\frac{cc}{w}\right]C+\dots+q_3=0\\
&\dots\dots\dots\dots\dots\dots
\end{aligned}\tag{64}$$

Upon inspection of these equations, there is seen to exist among them the same symmetry as shown, in Art. 44, among those for the adjustment of indirect observations. The diagonal coefficients, down to the right, are sums of squares and, therefore, always positive, and the other coefficients are symmetrical about this diagonal. However, the weights, w, occur here in the denominators of the coefficients, instead of in their numerators as in the previous case, and the constant terms, q, are the original closing errors of the condition equations, *in order.* The number of the normal equations will always be the same as that of the conditions, so the number of the q's will be the same.

The factors, A, B, C, etc., are obviously the same in number as the conditions, and they correspond to the various condition equations, *in order.* They are called *Correlates* or *Correlatives,* and are the unknowns of the normal equations, from which they are obtained by a solution according to the methods of the last chapter.

Substituting the values of the correlates, resulting from the solution of the normal equations, in the equations (63), we obtain the desired corrections, v_1, v_2, v_3, etc., which, applied to the corresponding observations, M_1, M_2, M_3, etc., give the best values, V_1, V_2, V_3, etc., of the observed quantities.

Thus, the process of adjustment may be stated in the *rule:* Write the condition equations involving the unknown corrections to the observations, and from them an equal number of normal equations, the solution of which gives the values of the correlates, from which the desired corrections to the observations are computed to obtain the best values of the observed quantities. The conditions (59) are first written, then the normal equations (64) are formed and solved, and lastly, the substitution of the correlates in (63) gives the desired corrections.

74. Observations of Equal Weight. By placing the weights equal to unity, in equations (63) and (64), we have the simpler forms:

$$\begin{aligned} v_1 &= a_1A + b_1B + c_1C + \dots \\ v_2 &= a_2A + b_2B + c_2C + \dots \\ &\dots\dots\dots\dots \\ v_n &= a_nA + b_nB + c_nC + \dots \end{aligned} \tag{65}$$

and the normal equations,

$$
\begin{aligned}
[aa]A+[ab]B+[ac]C+\ldots+q_1&=0\\
[ab]A+[bb]B+[bc]C+\ldots+q_2&=0 \qquad (66)\\
[ac]A+[bc]B+[cc]C+\ldots+q_3&=0
\end{aligned}
$$

Here the coefficients are the same and occur in the same order as those of the equations (22) in Art. 42, but the constant terms are simpler as they may be taken directly from the conditions without additional computation or combination.

75. Controls or Checks upon the Computation. The formation of the normal equations from the conditions is conveniently checked by means of sum-terms similar to those explained in Art. 43 for indirect observations. In this case, however, the sum-check does not include the constant terms, q, which are not formed in the same manner as the coefficients. Therefore, the check-equations have the form:

$$a_1+b_1+c_1+\ldots=s_1 \qquad (67)$$

which, multiplied by $\frac{a_1}{w_1}$, becomes:

$$\frac{a_1a_1}{w_1}+\frac{a_1b_1}{w_1}+\frac{a_1c_1}{w_1}+\ldots=\frac{a_1s_1}{w_1} \qquad (68)$$

Adding all such equations, we have,

$$\left[\frac{aa}{w}\right]+\left[\frac{ab}{w}\right]+\left[\frac{ac}{w}\right]+\ldots=\left[\frac{as}{w}\right] \qquad (69)$$

of which the left-hand member is the sum of the coefficients of the first normal equation. A similar equation can be written for each of the other normal equations. Thus, the sum-terms check the formation of the coefficients of the normal equations.

After the above checks have been verified, the constant term of each normal equation may be added to its sum-term, so as to form a check-term which shall include the constant, for use during the solution of the equations, as described in Art. 59.

The computation of the correlates is checked by substituting them in the normal equations other than the first. The corrections are checked by substitution in the condition equations.

And finally, the resulting values of the observed quantities may be substituted in the original conditions as a test of the correctness of the entire adjustment. This is the ultimate test of the work and should never be neglected. Beginners, in particular, should make use of all available checks.

76. Tabular Forms for Computations. By arranging the given data and the condition equations in the form of a table, the formation of the normal equations and the subsequent computation of the unknown corrections will be greatly facilitated. As the weights occur in the denominators of the coefficients, it is convenient to use their reciprocals throughout the computation. The following form is recommended:

Form for Condition Equations

(v)	v_1	v_2	v_3	v_4	v_5	v_6	v_7	Const.
$(1/w)$	$1/w_1$	$1/w_2$	$1/w_3$	$1/w_4$	$1/w_5$	$1/w_6$	$1/w_7$	(q)
(a)	a_1	a_2	a_3	a_4	a_5	a_6	a_7	q_1
(b)	b_1	b_2	b_3	b_4	b_5	b_6	b_7	q_2
(c)	c_1	c_2	c_3	c_4	c_5	c_6	c_7	q_3
......	......	......	......	......	......	......	......	......
(s)	s_1	s_2	s_3	s_4	s_5	s_6	s_7	

(70)

The parenthetical number at the left of each row is the symbol for the quantities in that row. The reciprocals of the weights are written in their proper columns just below the corresponding v's, while the sum of the coefficients is written at the foot of each column. It frequently happens that most of these coefficients, a, b, c, etc., are unity, so that the formation of the normal equation coefficients can be performed mentally, especially when the weights are equal.

The solution of the normal equations and the substitution back through the derived equations to obtain the correlates will be carried out by the abridged method of the last chapter in the form there given.

The computation of the corrections, also, may be tabulated conveniently, as follows:

COMPUTATION OF THE CORRECTIONS

	v_1	v_2	v_3	v_4	v_5	v_6	v_7
	$1/w_1$	$1/w_2$	$1/w_3$	$1/w_4$	$1/w_5$	$1/w_6$	$1/w_7$
(aA) (bB) (cC)	a_1A b_1B c_1C	a_2A b_2B c_2C	a_3A b_3B c_3C	a_4A b_4B c_4C	a_5A b_5B c_5C	a_6A b_6B c_6C	a_7A b_7B c_7C
....	......	......	......	......	......	......	......
(Sum)							
(Sum/w) (v)	v_1	v_2	v_3	v_4	v_5	v_6	v_7

(71)

The first row is obtained from the first row of (70) by multiplying the a's in succession by the first correlate, A; the second row, by multiplying the b's of (70) by the second correlate, B, etc. Multiplying each sum by the reciprocal of its weight gives the preliminary values of the v's, carried out, as were the correlates, one decimal place farther than is required in the final corrections, which are then written in the last row, taking the nearest figure in the next to the last place. These v's would be tested in the condition equations, and modified slightly, if necessary, so as to satisfy them exactly, the modifications being shown by canceling the changed figures and writing the adopted ones just to the right and above. Finally, the adopted values of the observed quantities would be tested in the original conditions, which they should rigidly satisfy.

77. Example: Adjustment of Levels. The following observed differences of elevation between the benchmarks, A, B, C, etc., will now be adjusted to the nearest hundredth, in accordance with the foregoing method. In Fig. 4, each line of levels is numbered, in parentheses, and an arrow shows the direction of running the levels over it, or rather, the direction of *stating* it, as it may have been run in either direction or both directions, but can be stated with only one sign which must correspond to a certain direction, plus or minus, according as the final benchmark is higher or lower than the initial one. Lines like that between C and G, which are parts of no complete circuit, do not enter into the adjustment in any way. The observed differences of elevation are as follows:

(1)	$+2.18$	(3)	-3.47	(5)	$+4.70$	(7)	-6.86
(2)	$+5.06$	(4)	$+1.32$	(6)	-9.82	(8)	$+3.46$

Let the weights of the lines (7) and (8) be 2 each, and those of the others, unity, or, for simplicity in the use of reciprocals, let the former be unity and the others, one-half, giving 1 and 2 for the reciprocals.

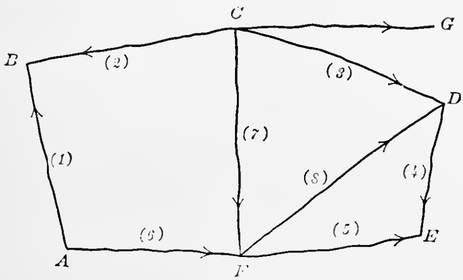


FIG. 4. System of Levels

As shown in Art. 71, each complete circuit furnishes the condition that the sum of its adjusted differences of elevation shall equal zero when given the proper signs as if run continuously around the circuit, clockwise or counter-clockwise. Also, the number of *independent conditions* is the same as the number of extra observations above those necessary to connect the given benchmarks. These necessary lines may be drawn, one at a time, starting at one benchmark, as long as a new benchmark is added for each line drawn. Then each line added to the figure, between two benchmarks already shown, gives one independent condition which should always be written so as to *include that line*. When the complete figure has been reproduced on paper, in this manner, omitting no lines, all of the necessary, independent conditions for the adjustment will be indicated. Their number may be verified by the rule that it is the same as the number of extra observations. Also from the above construction,

Number of conditions = (Number of lines)
− (Number of bench-marks) + 1.

Assuming the lines (1), (2), (3), (4), and (5) to be those necessary to connect all of the benchmarks, we write a condition for each of the remaining lines, namely, (6), (7), and (8). It is essential, of course, that all of the lines which form circuits should appear in the conditions. Thus we obtain the original conditions,

$$\begin{aligned} +V_1 - V_2 + V_3 + V_4 - V_5 - V_6 &= 0 \\ +V_1 - V_2 + V_7 - V_6 &= 0 \\ +V_5 - V_4 - V_8 &= 0 \end{aligned} \qquad (72)$$

The minus signs result from changing the directions of the arrows so as to be continuous around each circuit. It is not necessary that all the circuits be traversed in the same direction, however, in a given problem. Substituting for each V its observed value, we find the closure errors of these three circuits to be, respectively, +0.09, +0.08, and −0.08. Therefore, the following table may be formed directly, the coefficients being unity:

Condition Equations

(r)	v_1	v_2	v_3	v_4	v_5	v_6	v_7	v_8	Const.
$(1/w)$	2	2	2	2	2	2	1	1	(q)
(a)	+1	−1	+1	+1	−1	−1			+0.09
(b)	+1	−1				−1	+1		+0.08
(c)				−1	+1			−1	−0.08
(s)	+2	−2	+1	0	0	−2	+1	−1	

(73)

The normal equations may be written by inspection, owing to the simplicity of the condition equations. Each square or product is formed in the proper column, above, and multiplied by its $1/w$; the sum of all similar results being the coefficient for the normal equation. Thus, there are six aa's, each being $+1$, three ab's, each being $+1$, and two ac's, each of which is -1. Where no coefficient appears in the condition equation for a certain v, that coefficient is regarded as zero. Each aa/w will be $2\times(+1)=+2$, as also, will be each ab/w. As there is no column in which both b and c occur, the products, bc, are zero. When all the coefficients in any condition equation are unity, the sum of the squares, each divided by its weight, is equal to the sum of the reciprocals of the weights; thus, $[bb/w]=2+2+2+1=+7$. Likewise, the product terms may be written by inspection, but the signs must be carefully considered; thus, $[ac/w]=-2-2=-4$, etc. The sum-terms are treated in the same way as the coefficients, to test the correctness of the computations.

NORMAL EQUATIONS

A	B	C	(q)	Sum
+12	+6	−4	+0.09	+14
	+7	0	+0.08	+13
		+5	−0.08	+ 1

(74)

It must be remembered that the sum includes all the coefficients of an equation, whether written or not, so that, when the abridged form is used, as above, the coefficients must be read down and to the right as explained in Art. 57 (c).

Preparatory to solving the normal equations, the constants are added to their respective sum-terms to form the check-terms for use throughout the solution, in order that the operations performed upon the constants may be included in the checks. In their form for solution, therefore, these equations are:

NORMAL EQUATIONS

A	B	C	Const.	Check
+12	+6	−4	+0.09	+14.09
	+7	0	+0.08	+13.08
		+5	−0.08	+ 0.92

(75)

These equations will now be solved by the Abridged Method as explained in Art. 60, the separate operations being indicated.

	A	B	C	Const.	Check	
(I)	+12	+6	−4	+0.09	+14.09	
(2)		+7	0	+0.08	+13.08	
(3)			+5	−0.08	+ 0.92	
(2)		+7	0	+0.08	+13.08	
(4)		−3	+2	−0.04	− 7.04	(I)×(−6/12)
(II)		+4	+2	+0.04	+ 6.04√	(2)+(4)
(3)			+5	−0.08	+ 0.92	
(5)			−1.33	+0.03	+ 4.70	(I)×(+4/12)
(6)			−1.00	−0.02	− 3.02	(II)×(−2/4)
(III)			+2.67	−0.07	+ 2.60√	(3)+(5)+(6)

Correlates

	A	B	C
Constants	−0.09	−0.04	$\dfrac{+0.07}{+2.67}=$
C-terms	+0.104	−0.052	$+0.026=C$
B-terms	+0.138	$\dfrac{-0.092}{+4}=$	
	$\dfrac{+0.152}{+12}=$	$-0.023=B$	
	$+0.013=A$		

(76)

Tests of Correlates

Equation (2) $+0.078-0.161+0.08=-0.003$

Equation (3) $-0.052+0.130-0.08=-0.002$

These discrepancies would be reduced by carrying out the corre-

CORRECTIONS

	v_1	v_2	v_3	v_4	v_5	v_6	v_7	v_8
$(1/w)$	2	2	2	2	2	2	1	1
(aA)	+0.013	−0.013	+0.013	+0.013	−0.013	−0.013		
(bB)	−0.023	+0.023				+0.023	−0.023	
(cC)				−0.026	+0.026			−0.026
(Sum)	−0.010	+0.010	+0.013	−0.013	+0.013	+0.010	−0.023	−0.026
(Sum/w)	−0.020	+0.020	+0.026	−0.026	+0.026	+0.020	−0.023	−0.026
(v)	−0.02	+0.02	+0.03	−0.03	+0.03	+0.02	−0.02	$-0.0\cancel{3}^{2}$

TESTS OF CORRECTIONS

1st Condition: −0.02 −0.02 +0.03 −0.03 −0.03 −0.02 +0.09 = 0.00✓

2d Condition: −0.02 −0.02 −0.02 −0.02 +0.08 = 0.00✓ (78)

3d Condition: +0.03 +0.03 $+0.\cancel{3}^{2}$ $-0.08 = 0.\cancel{1}^{0}$

lates to four places, instead of three. But these are satisfactory when the corrections are desired to hundredths, only.

Upon testing the corrections by substituting them in the condition equations (73), it is found that the third condition fails by +0.01. This discrepancy must be removed by arbitrarily altering one or more of the corrections involved in it, as it is due to neglected remainders. At the same time, the other two conditions, which check exactly, must not be disturbed. Therefore, it is desirable to find a correction which is used in the third condition only. Such a one is v_8 which is seen to be too large by 0.004, and which, moreover, belongs to one of the observations of greater weight so that it would be expected to have a smaller correction. There is reason, therefore, for reducing this correction by the necessary 0.01 in order to satisfy the condition. The change is made as indicated so that the original figure remains. If there were no single correction which could be modified without affecting other conditions, it might be necessary to alter two or three corrections in order to satisfy all the conditions by a given set of corrections.

The final test of the correctness of the adjustment consists in substituting the adjusted values of the differences of elevation in the original conditions, (72), or in the other conditions which were not used because not independent of these. It is well to restate the conditions, using the corrected differences of elevation, in order to secure a check on the condition equations. Referring to the diagram, Fig. 4, therefore, and applying to each observation the corresponding correction, we have:

FINAL TESTS OF THE ADJUSTED VALUES

$$
\begin{aligned}
&\text{Circuit } A\text{–}B\text{–}C\text{–}F: && +2.16-5.08-6.88+9.80=0.00\checkmark \\
&\text{Circuit } C\text{–}D\text{–}F: && -3.44-3.44+6.88 \qquad =0.00\checkmark \qquad (79)\\
&\text{Circuit } D\text{–}E\text{–}F: && +1.29-4.73+3.44 \qquad =0.00\checkmark
\end{aligned}
$$

These comprise all the elemental circuits, so that any combinations of these would also be satisfied.

78. Arrangement of Equations. The larger the coefficients of the normal equations, the greater will be the labor of solution, generally speaking, so it is important, as was shown in the case of indirect observations, to make them as small as practicable. The methods of Arts. 48 and 49 do not apply directly to conditioned observations, but it is possible to *select* the conditions and *arrange* the condition equations in such a manner as to save some labor in the solution of the normal equations.

Inspection of the equations (73) and (74) shows that the shorter conditions, that is, those which involve fewer observations, will produce smaller coefficients for the normal equations. Therefore, it is important to select the shorter conditions, as far as practicable. In the above example, for instance, the three small circuits might have been used to advantage, although, in so short a problem the advantage is less evident than in longer ones.

It is apparent, also, that by arranging the condition equations in a certain order, with the shorter ones first, the larger coefficients will occur later in the normal equations, instead of earlier, which is an advantage especially in the abridged method of solution. Then, too, it is possible to place those equations first which have no terms in common, so that the product-terms, $[ab/w]$, $[ac/w]$, etc., in the first normal equation, may be zero in some cases. Each of such zero coefficients gives a zero elimination-factor which saves writing a whole line in the elimination. In some problems this is very important. In the above example, if the second and third equations had been written as the first and second, respectively, $[ab/w]$ would have been zero, thus saving the second step in the elimination, since the second normal equation would have had no A-term and so would have been, itself, the first derived equation, number (II). Sometimes, it is possible to save several steps in the elimination in this manner.

79. Example: Local Adjustment of Angles by the Method of Correlates. In triangulation, the methods of measuring the angles at a station may result in several *extra* angles being observed. As shown in the latter part of Art. 71, each of these extra observations yields one independent condition. To illustrate the method of adjusting the angles so as to satisfy all these conditions, we shall consider the case shown in that article, Fig. 3, assuming the weights to be equal.

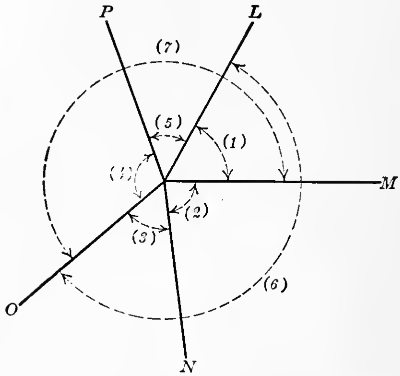


FIG. 3. Horizontal Angles at a Station

OBSERVED ANGLES

$$
\begin{array}{llll}
M_1 = 85° \; 14' \; 24.5'' & & M_5 = \;\; 50° \; 23' \; 26.7'' & \\
M_2 = 83 \;\; 45 \;\; 32.0 & & M_6 = 210 \;\; 35 \;\; 17.5 & \qquad (80) \\
M_3 = 41 \;\; 35 \;\; 24.0 & & M_7 = 234 \;\; 39 \;\; 08.2 & \\
M_4 = 99 \;\; 01 \;\; 14.1 & & &
\end{array}
$$

Adopting (1), (2), (5), and (6), as the necessary angles, the conditions may be written,

CONDITION EQUATIONS

$$
\begin{aligned}
V_1+V_2+V_3-V_6 \quad &= 0 \\
V_4+V_5+V_6-360° &= 0 \qquad (81) \\
-V_1+V_6+V_7-360° &= 0
\end{aligned}
$$

from which, by substituting for each V its value, $M+v$, we have,

REDUCED CONDITION EQUATIONS

	v_1	v_2	v_3	v_4	v_5	v_6	v_7	
(a)	+1	+1	+1			−1		+3.0=0
(b)				+1	+1	+1		−1.7
(c)	−1					+1	+1	+1.2
(s)	0	+1	+1	+1	+1	+1	+1	

(82)

NORMAL EQUATIONS

A	B	C	Const.	Sum	Check
+4	−1	−2	+3.0	+1✓	+4.0
	+3	+1	−1.7	+3✓	+1.3
		+3	+1.2	+2✓	+3.2

(83)

Solving these equations as in Art. 77, we obtain the following values of the correlates:

$$A = -1.347 \qquad B = +0.619 \qquad C = -1.503$$

whence the corrections to the observed angles are.

v_1	v_2	v_3	v_4	v_5	v_6	v_7
+0.16″	−1.35″	−1.35″	+0.62″	+0.62″	+0.46″	−1.50″

(84)

which exactly satisfy the given conditions.

The adjusted values of the angles, therefore, are,

$$\begin{aligned}
V_1 = M_1 + v_1 &= 85° \; 14' \; 24.5'' + 0.16'' = 85° \; 14' \; 24.66'' \\
V_2 &= 83 \; 45 \; 32.0 - 1.35 = 83 \; 45 \; 30.65 \\
V_3 &= 41 \; 35 \; 24.0 - 1.35 = 41 \; 35 \; 22.65 \\
V_4 &= 99 \; 01 \; 14.1 + 0.62 = 99 \; 01 \; 14.72 \\
V_5 &= 50 \; 23 \; 26.7 + 0.62 = 50 \; 23 \; 27.32 \\
V_6 &= 210 \; 35 \; 17.5 + 0.46 = 210 \; 35 \; 17.96 \\
V_7 &= 234 \; 39 \; 08.2 - 1.50 = 234 \; 39 \; 06.70
\end{aligned} \tag{85}$$

As a final test of the adjustment, these adjusted values are substituted in the original conditions, which they are found to satisfy.

80. Special Case of One Condition Only. The general condition equation is the first one of (59), namely,

$$a_1v_1 + a_2v_2 + a_3v_3 + \; . \; . \; . \; + a_nv_n + q = 0 \tag{86}$$

whence, the single normal equation is

$$[aa \cdot w]A + q = 0 \tag{87}$$

and the corrections, from (63), are

$$v_1=(a_1/w_1)A; \quad v_2=(a_2/w_2)A; \quad \text{etc.} \tag{88}$$

The case of special interest, however, is that in which the coefficients of the condition equation are unity; thus,

$$v_1+v_2+v_3+\ .\ .\ .\ +v_n+q=0 \tag{89}$$

The normal equation, then, becomes,

$$[1/w]A+q=0 \tag{90}$$

so that

$$A=-q/[1/w] \tag{91}$$

The corrections, with this value of A, are, therefore,

$$v_1=-q\frac{1/w_1}{[1/w]}; \quad v_2=-q\frac{1/w_2}{[1/w]}; \quad \text{etc.} \tag{92}$$

Thus the corrections are proportional to the reciprocals of the weights, and each correction is equal to the total closure correction divided by the algebraic sum of the reciprocals of the weights and multiplied by the reciprocal of the corresponding weight.

For example, suppose we have a single circuit of levels which add to $+0.24$ instead of zero, and that the weights of the nine lines are 2, 3, 1, 2, 3, 1, 1, 3, and 1. The least common multiple of the weights is 6, and they may be written, 2/6, 3/6, 1/6, 2/6, 3/6, 1/6, 1/6, 3/6, and 1/6, respectively, so that their reciprocals are the following integers, in order, 3, 2, 6, 3, 2, 6, 6, 2, and 6, whose sum is 36. The corrections, therefore, are obtained by multiplying each of these reciprocals into the constant, $-0.24/36$, resulting thus:

$$\begin{array}{lllll} -0.020, & -0.013, & -0.040, & -0.020, & -0013, \\ -0.040, & -0.040, & -0.013, & -0.040 & \end{array}$$

Testing these corrections in (89), their sum is -0.239 instead of -0.24, so that it is necessary to add -0.001 to one of them, preferably changing -0.013 to -0.014, in order to rigidly satisfy the prescribed condition.

The important point is that the corrections may be written by inspection, in such cases, from the fact that they are proportional to the reciprocals of the weights and that their sum must be equal to $-q$. If any of the v's in the condition equation be *negative*, the signs of the corresponding corrections are changed. Thus, if the condition equation were

$$v_1 - v_2 + v_3 - v_4 + v_5 - v_6 + v_7 + v_8 + v_9 + 0.24 = 0$$

the corrections would be numerically the same as above, but the signs of the second, fourth, and sixth would be plus instead of minus. In testing the corrections in the condition equation, then, these three would be multiplied by -1, so that the condition would be satisfied as before.

This method of distributing the error of closure is somewhat similar to that used in the special case of weighted mean of two quantities, given in Art. 36.

81. Adjustment by the Method of Indirect Observations. It is possible to adjust conditioned observations as if the quantities observed were independent, that is, by the method used in Chap. III for indirect observations. Although this process is generally longer and less satisfactory than the solution by the method of correlates, it will be explained, briefly, in order that it may be used when the circumstances are favorable, and that the subject may be better understood.

In Art. 71 it was shown that a certain number of observations would be *necessary*, in a given problem, for the determination of the unknown quantities, on the assumption that those observations were correct, and that the remaining, *extra*, observations would furnish one condition each, to be satisfied by the adjusted quantities. Let those observations which are selected as the necessary ones be stated simply as observation equations, namely,

$$V_1 = M_1; \quad V_2 = M_2; \quad V_3 = M_3; \quad \text{etc.} \tag{93}$$

Then each condition, selected so as to involve but one new quantity, may be expressed in terms of the other quantities only, so that the total number of unknowns will not exceed those first, necessary ones. From the entire set of these observation equations, the normal equations are formed, as many as there are necessary (i.e., independent) unknowns, and their solution gives the adjusted values of the quantities.

For example, in the local adjustment of angles at a station, in Art. 79, and Fig. 3, the three conditions could be replaced by observation equations, as follows:

Conditions	Observation Equations	
$V_1+V_2+V_3-V_6 \quad =0$	$-V_1-V_2+V_6 \quad =M_3$	
$V_4+V_5+V_6-360°=0$	$-V_5-V_6+360°=M_4$	(94)
$-V_1+V_6+V_7-360°=0$	$+V_1-V_6+360°=M_7$	

The entire seven observation equations, therefore, are,

$$
\begin{array}{llllllll}
+V_1 & & & & -M_1 & & =0 & \text{Wt.}=w_1 \\
& +V_2 & & & -M_2 & & =0 & w_2 \\
-V_1 & -V_2 & & +V_6 & -M_3 & & =0 & w_3 \\
& & -V_5 & -V_6 & -M_4 & +360° & =0 & w_4 \\
& & +V_5 & & -M_5 & & =0 & w_5 \\
& & & +V_6 & -M_6 & & =0 & w_6 \\
+V_1 & & & -V_6 & -M_7 & +360° & =0 & w_7
\end{array}
\qquad (95)
$$

in which there are but the four unknowns, namely, the angles, (1), (2), (5), and (6), and each M represents an observed value or constant term. These equations correspond to (16). Forming and solving the four normal equations by the methods of Chap. III, the best values of the angles are determined directly.

82. Example: Local Adjustment of Angles as Independent Quantities. The solution of the above example will be continued, to illustrate the method, but with equal weights, for simplicity. Let the observed angles be the same as those used in Art. 79, as the comparison of the two methods will be useful. These angles are given in (80). Substituting their values in (95), and also for each V, the corresponding $M+v$, so as to reduce the constant terms, which is equivalent to assuming for these V's the corresponding M's as approximate values, as in Art. 48, we have the simplified observation equations:

v_1	v_2	v_5	v_6	(l)		(s)
$+1$					$=0$	$+1$
	$+1$				$=0$	$+1$
-1	-1		$+1$	$-3.0''$	$=0$	-4.0
		-1	-1	$+1.7$	$=0$	-0.3
		$+1$			$=0$	$+1$
			$+1$		$=0$	$+1$
$+1$			-1	-1.2	$=0$	-1.2

(96)

The normal equations are obtained by inspection;

v_1	v_2	v_5	v_6	Const.		Sum
$+3$	$+1$		-2	$+1.8$	$=0$	$+3.8$ ✓
	$+2$		-1	$+3.0$	$=0$	$+5.0$ ✓
		$+2$	$+1$	-1.7	$=0$	$+1.3$ ✓
			$+4$	-3.5	$=0$	-1.5 ✓

(97)

Solving them, we obtain directly,

$$v_1 = +0.17'' \quad v_2 = -1.36'' \quad v_5 = +0.61'' \quad \text{and} \quad v_6 = +0.48'' \tag{98}$$

whence,

$$\begin{aligned} V_1 &= 85^\circ\ 14'\ 24.67'' & V_5 &= \ \ 50^\circ\ 23'\ 27.31' \\ V_2 &= 83\ \ \ 45'\ 30.64 & V_6 &= 210\ \ \ 35\ \ \ 17.98 \end{aligned} \tag{99}$$

by a combination of which, the angles V_3, V_4, and V_7, are computed. Thus,

$$\begin{aligned} V_3 &= V_6 \quad - V_1 - V_2 = \ \ 41^\circ\ 35'\ 22.67'' \\ V_4 &= 360^\circ - V_5 - V_6 = \ \ 99^\circ\ 01'\ 14.71'' \\ V_7 &= 360^\circ - V_6 + V_1 = 234^\circ\ 39'\ 06.69'' \end{aligned} \tag{100}$$

This completes the adjustment. A comparison of these results with those of (85) in Art. 79, shows them to be only slightly different.

83. Comparison of the Two Methods. The principal points of difference between the methods lie in the labor involved and in the checks which are available. It may have been noted that in the preceding article there is no ultimate check upon the corrections or the adjusted angles. The corrections must satisfy the normal equations, of course, as in any other adjustment, but there is no check upon the observation equations (96). The checks afforded by the conditions, in the method of correlates, are forfeited in the method of indirect observations, being used for the determination of some of the unknowns, as V_3, V_4, and V_7 in the above problem. This is an evident disadvantage of the latter method, inasmuch as the final check is very desirable and important. The sum-checks controlling the formation and solution of the normal equations are present in both methods.

In the method of correlates, the number of normal equations is equal to the number of conditions, which must be less than that of the unknown quantities or observations. In the method of indirect observations on the other hand, the number of normal equations is that of the necessary, independent unknowns, and therefore may be greater or less than in the former method. Usually, however, the number of conditions is small as compared with the number of independent unknowns, so that the method of correlates is likely to be the shorter, although the determination of the corrections from the correlates is a step which is not required in the other method where the unknowns are obtained directly from the solution of the normal equations, or at most, by a single addition or multiplication. If the number of conditions happens to be nearly as great as that of the independent unknowns, as in the above example, the disadvantages of the method of indirect observations are less, and the simplicity of the normal equations, resulting from the considerable number of zeros in the observation equations, may give this method the advantage, even, although this is seldom likely to be the case. Moreover, the absence of the final check in the conditions is a serious defect, and gives to the method of correlates the preference.

84. Adjustments not Rigid. The final, adjusted values of the unknown quantities cannot be regarded as the *correct* ones, of course, but are approximations to them. As different methods may be used in the adjustment, and as different sets of conditions may be used in the same method, it is obvious that small discrepancies are likely to exist between the final values obtained from different adjustments of the same data. Each of these sets of results may satisfy all of the conditions as required and may constitute an adjustment which is entirely satisfactory. Usually, the discrepancies will be so small as to be negligible as compared with the accidental errors of the observations.

CHAPTER VI

ADJUSTMENT OF TRIANGULATION

85. Triangulation. A system or network of triangulation consists of a series of stations connected by lines in such a manner as to form triangles having their vertices at the stations. The length of one line, called the base-line, being determined by direct measurement, usually with a tape, and the horizontal angles between the lines at each station being measured with a transit or theodolite, the lengths of all the lines become known by computation from the base-line and the angles through successive triangles. The differences of elevation between the stations are obtained from observed vertical angles which determine the elevations of the stations above sea-level when one of them has been connected to sea-level by a line of precise, or geodetic, leveling. The position of the system on the earth's surface is fixed by astronomical observations for the latitude of one station, the longitude of one station, and the azimuth of one line. The *size* of the system, or net, depends, therefore, upon the length of the base; its *shape*, or form, depends upon the horizontal angles; its *position*, upon the astronomical observations; and its *elevations*, upon the vertical angles and the initial elevation. If the triangulation be based upon, or connected to, two stations of another system which has been completely determined and adopted in size, position, and elevation, the line joining the two stations may be used as the base-line for the new work, and the azimuth and the latitude, longitude, and elevation of one of the stations will determine the position and initial elevation of the new net. However, if the new triangulation be complete in itself in regard to one or more of these elements, and in addition be connected to previously adjusted and adopted work, this connection affords checks upon the corresponding elements, and therefore, from one to five conditions must be satisfied if all of the work is to be made consistent as to length, latitude, longitude, azimuth, and elevation. The shape of the net, and the differences of elevation, therefore, must be adjusted so as to fulfill these requirements. Moreover, the horizontal angles must be adjusted to conform to certain geometrical and trigonometrical conditions which depend upon the arrangement of the lines and stations and the angles observed.

The vertical angles are independent of the horizontal ones and are adjusted by themselves in any case. The adjustment of a system so as to close upon fixed, or adopted, work with regard to any of the five elements of length, latitude, longitude, azimuth, and elevation will be discussed farther on.[1] There remains, then, the adjustment of a system which is complete in itself. In this, the length of the base and the initial latitude, longitude, azimuth, and elevation are determined separately and independently of the horizontal angles in the net, and so do not enter into the adjustment as long as there is but one of each of these elements. The adjustment of the horizontal angles, therefore, will now be considered.

86. Nature of the Conditions. The horizontal angles in triangulation are subject to two classes of geometrical conditions, namely, those which involve the angles at one station only, and those which define relations between the angles at two or more stations. The former are called local conditions and the latter figure conditions, giving rise to local and figure adjustments.

The *local conditions* express the requirement that the adjusted values of the observed angles at a station shall satisfy the indicated horizon-closures and algebraic sums.

The *figure conditions* are of two kinds, known as angle equations and side equations. An angle equation requires that the sum of the angles of a triangle or polygon shall be equal to the number of right angles prescribed by geometry for a plane figure, plus the spherical excess. A side equation requires that if the

[1] Art. 106, et seq.

length of a line in the figure be computed from another line through two different series of triangles, that is, by two different routes, the two results must be equal.

Since all of these conditions must be satisfied simultaneously, they would enter into a single adjustment, ordinarily. As will be explained later, however, it may be convenient to perform the local adjustment separately, prior to the figure adjustment, the latter being so arranged as not to disturb the former.

87. Local Adjustment. In modern field practice, simplicity is sought for the sake of economy. Accordingly, observations are arranged, as far as practicable, so as to lessen the office work necessary for their reduction, but without a sacrifice of precision. The angles at a station, therefore, are observed in such a manner as to avoid combinations which introduce checks and conditions requiring extensive local adjustment. It is customary to measure one angle for each of the signals less one, and then a single one to close the horizon, thus securing one check which involves all of the observed angles. The local adjustment is thereby reduced to one simple condition, with equal weights, also, in most cases, so that it amounts to a mere distribution of the error of closure, as explained in Art. 80. If extra observations have been made, however, so that two or more conditions are to be satisfied, the general method of adjustment must be used. This has been demonstrated in Arts. 79 and 82, in the last chapter, in which the number of the conditions was shown to be equal to the number of extra observations. Thus, if S stations be observed, $S-1$ angles would be sufficient to connect them, and if N angles be measured between them, the number of extra observations, and therefore, the number of local conditions may be expressed in the formula,

$$\text{Number of Local Conditions at a Station} = N - S + 1 \qquad (101)$$

88. Figure Adjustment. Notation. In order to distinguish between stations occupied and unoccupied, and lines observed in both directions or in one direction only, lines shown in diagrams of triangulation will be broken at the ends from which they are not observed, full lines indicating observation at both ends. Stations which are sighted upon but not occupied with the instrument will be recognized from the fact that all the lines at those stations will be broken. Thus, in Fig. 5, the station *Pan* was

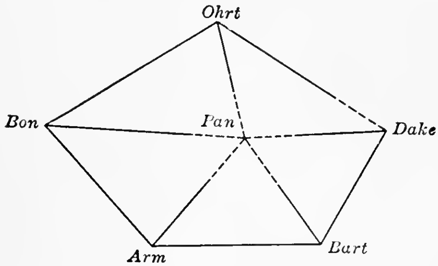


FIG. 5. Unobserved Lines and Unoccupied Station

not occupied, as no full lines radiate from it. *Dake* was occupied and *Pan* and *Bart* were observed from it, but *Ohrt* was not observed from it, although *Dake* was observed from *Ohrt*. The other stations were occupied completely as shown by the lines being unbroken at those ends.

89. Classification of Figures. Although the figures in triangulation may be very complicated and the adjustment very laborious, the work in such a case loses its economic advantages of covering a great area or distance at the minimum of cost consistent with the accuracy desired. In the best practice, therefore, simple figures are used, and special attention is given to measuring each angle with the requisite degree of precision. These simple figures may be classified as *triangles*, *quadrilaterals*, and *central-point figures*. A triangle consists of three stations connected by three lines. A quadrilateral has four stations connected by six lines. A central-point figure is a polygon with a station at each vertex and another station in the interior from which lines radiate to the vertices; the polygon usually has not more than six sides. The lines in these figures may be full or partly broken, as above. Fig. 5 represents a central-point figure. A typical quadrilateral with diagonals is shown in Fig. 6, while Fig. 7 is the simplest form of a central-point figure, which may be considered, also, as a quadrilateral. In Fig. 8 is shown a combination of a central-point figure with a polygon having diagonals;

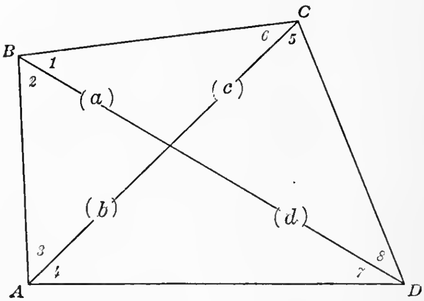


FIG. 6. Quadrilateral

this is seen to increase the intricacy of the system, which would have been a simple central-point figure had the diagonals KM and MO been omitted.[1]

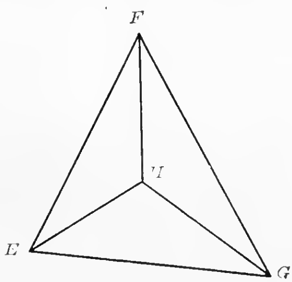


FIG. 7. Central Point Figure

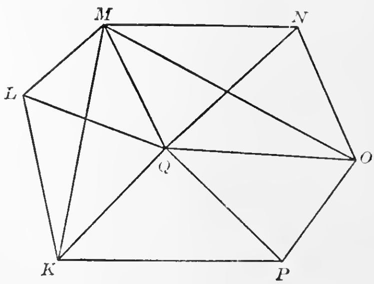


FIG. 8. Central Point Figure with Extra Diagonals

90. Angle Equations. The triangle is the unit figure in triangulation. For each triangle or other polygon of which all the angles have been observed, an angle equation may be written expressing the condition that the sum of the adjusted angles shall

[1] In the diagrams representing triangulation, it is assumed that there is no station at the intersection of diagonals of a figure unless there is an angle at that point in one of them. If, in the remote case, a station happened to fall at this intersection, the diagram would be slightly distorted so as to indicate the fact without question.

be the theoretical amount, namely, a certain number of right angles plus the spherical excess, ϵ, of the figure.[1] Thus, referring to Fig. 6, in which the separate angles are numbered clockwise at each station, and representing their adjusted values by V's, as usual, the four triangles yield the following angle equations,

$$
\begin{array}{llll}
\text{Triangle } (a) & ABC, & V_1+V_2+V_3+V_6-(180°+\epsilon_a)=0 & \\
\quad " \quad (b) & DAB, & V_2+V_3+V_4+V_7-(180°+\epsilon_b)=0 & (102) \\
\quad " \quad (c) & DBC, & V_1+V_5+V_6+V_8-(180°+\epsilon_c)=0 & \\
\quad " \quad (d) & DAC, & V_4+V_5+V_7+V_8-(180°+\epsilon_d)=0 &
\end{array}
$$

in which *a, b, c,* and *d* refer to the separate triangles as shown in the figure. Since the spherical excess depends directly upon the area of a figure,[2] the excess for the entire quadrilateral should be equal to the sum of the two excesses for the pair of triangles formed by each diagonal. Therefore,

$$\epsilon_a+\epsilon_d=\epsilon_b+\epsilon_c \tag{103}$$

which affords a check upon their computation. By inspection, then, we find that from any three of the above angle equations it is possible to derive the fourth by addition and subtraction. Also, from the whole quadrilateral, we may write the condition,

$$V_1+V_2+V_3+V_4+V_5+V_6+V_7+V_8-(360°+\epsilon_a+\epsilon_d)=0 \tag{104}$$

and this equation is seen to be the sum of the first and the fourth of (101). Therefore, any three of the four triangles may be selected from which to write the three independent conditions or angle equations. In other words, if two triangles formed by a single diagonal satisfy their conditions, the entire figure must satisfy its condition (104); then if a third triangle condition, also, be satisfied, the fourth one is sure to be, since the fourth triangle is equal to the whole figure minus the third one.

If we adopt the first three of the equations (102) as the independent ones, and write for each V, in the usual manner, its value,

[1] It is seldom that an angle equation has to be written for a figure greater than a triangle, as an open quadrilateral (without a diagonal) is not rigid and should be avoided.

[2] From spherical trigonometry.

$M+v$, in which M is the observed value of the angle and v is its correction to be obtained from the adjustment, we have for this quadrilateral the following set of final angle equations, corresponding to (59):

$$\begin{aligned} v_1+v_2+v_3+v_6+q_1&=0 \\ v_2+v_3+v_4+v_7+q_2&=0 \\ v_1+v_5+v_6+v_8+q_3&=0 \end{aligned} \tag{105}$$

in which q is the error of closure of a triangle, positive when the sum of the observed angles is too large. These closures may be checked in the same manner as the spherical excesses, that is, the closure for the whole figure must be the sum of the closures for each pair of component triangles. Thus,

$$q_a+q_d=q_b+q_c \tag{106}$$

91. Number of Angle Equations in a Figure. To determine the number of independent angle equations in a given figure, A–B–C–D–E, Fig. 11, we may proceed as follows: Start with

FIG. 9.

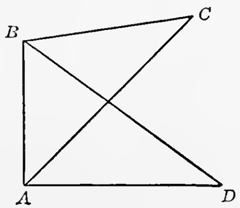


FIG. 10

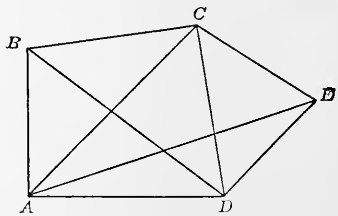


FIG. 11.

Determination of the Number of Angle Equations in a Figure

two stations, A and B, connected by one line, as in Fig. 9. Add the station C, with two lines to A and B, and one triangle is obtained. Add station D and *two* lines, to A and B, and a second triangle is formed, as in Fig. 10. Add the *third* line, from D to C, and the quadrilateral is completed, making three independent angle equations in all. If another station, E, be added, with *three* lines to A, C, and D, as in Fig. 11, two of these lines form a new triangle, as before, and the third completes a quadrilateral, A–C–E–D, in which one triangle, A–C–D, formed a part of the previous figure, A–B–C–D, and is therefore already included in the conditions. The third line from E thus adds but one new condition, making a total of five angle equations. If the line BE were added, it would be the second diagonal in the quadrilateral A–B–C–E, of which two triangles are already included in the figure, so that the third angle equation, only, for that quadrilateral would be added by this line.

We may generalize from this procedure and write a formula for the number of independent angle equations in any figure. Starting with three lines and three stations in the form of a triangle, we have one condition. If we add one station and one line to it, no new conditions are introduced, but each additional line to that station gives one new condition, and the same is true of further additions of stations and lines until the entire figure has been drawn. Therefore, each station added to the initial triangle adds as many conditions as there are lines, less one, running to that station, so that the total number of conditions would be the aggregate of these conditions together with the one for the initial triangle, that is, the whole number of lines minus the whole number of stations, plus one. But if any one of these lines be unobserved at one end, one angle of the corresponding triangle would be missing and that line would not count for a condition. Also, if one station were entirely unoccupied, as *Pan*, in Fig. 5, page 83, it could enter into no complete triangle and would have to be omitted from the stations counted in determining the number of angle equations. Finally, then, we may write the following formula for the number of independent angle equations in a given figure:

$$\text{Number of Angle Equations} = L' - S' + 1 \qquad (107)$$

in which L' is the number of *full lines* in the figure and S' is the number of *occupied stations*.

92. Side Equations. The triangles of a figure may close exactly to $180° + \epsilon$, and the angles still be inconsistent with regard to the closure of the whole figure when the lengths are computed. To illustrate this, suppose the triangles of Fig. 6, page 84, to be plotted in the following order, the angles of each having been adjusted to a closure and the local conditions satisfied. Plot triangle (a) to any convenient scale, as in Fig. 12, using the given angles. Upon the side AB, construct triangle (b) with vertex falling at d'. Upon BC, construct triangle (c) and its vertex might fall at d''. Then if triangle (d) were plotted upon AC as a base, its vertex might fall at d''', and the angle at d''' would still equal the sum of d' and d'' as required by the local condition at station D. Thus, the lengths of the lines running to this station would fail to check because they did not intersect at a single point. The side equation for this figure, therefore, would require

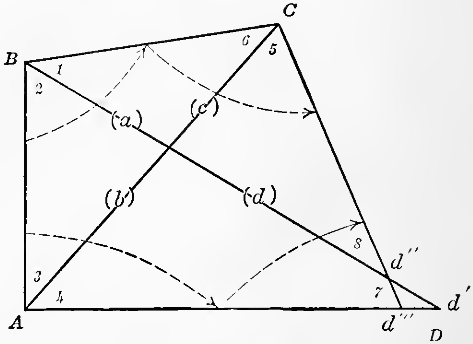


Fig. 12. Side Equation

that the three points, d', d'', and d''', be coincident, which is equivalent to the requirement that d'' and d''' be coincident, or that the line Cd'' in the triangle (c) be equal to the line Cd''' in the triangle (d). In other words, the side equation requires that if one line of a figure be computed from another line through two series of triangles, or by two different routes, the resulting lengths shall be equal. Since the same initial line is used in both cases, it cancels from the equation, and the angles, only, are concerned in the condition.

93. Side Equation of a Quadrilateral. Starting with the line AB and computing CD through the triangles (a) and (c), as indicated by the upper dotted arrows, and using the final, adjusted values of the angles, which make the points d', d'', and d''' coincident, we have,

$$CD = BC \cdot \frac{\sin V_1}{\sin V_8} = AB \frac{\sin V_3 \sin V_1}{\sin V_6 \sin V_8} \qquad (108)$$

Likewise, computing through the triangles (b) and (d),

$$CD = AD\frac{\sin V_4}{\sin V_5} = AB\frac{\sin V_2}{\sin V_7}\frac{\sin V_4}{\sin V_5} \tag{109}$$

Equating these expressions for CD and canceling the factor AB,

$$\frac{\sin V_3}{\sin V_6}\frac{\sin V_1}{\sin V_8} = \frac{\sin V_2}{\sin V_7}\frac{\sin V_4}{\sin V_5} \tag{110}$$

Multiplying both members of (110) by the reciprocal of the second member, we obtain a statement of the side equation in the form

$$\frac{\sin V_1}{\sin V_2}\frac{\sin V_3}{\sin V_4}\frac{\sin V_5}{\sin V_6}\frac{\sin V_7}{\sin V_8} = 1 \tag{111}$$

in which the numerator contains the odd-numbered angles and the denominator, the even ones, which happens as a result of our numbering the two angles at each station in clockwise order, and which is a useful check on the formation of the side equation. To reduce this equation to the first degree, we take the logarithm of each member and equate them, whence,

$$\begin{aligned}\text{logsin } V_1 + \text{logsin } V_3 + \text{logsin } V_5 + \text{logsin } V_7 \\ - \text{logsin } V_2 - \text{logsin } V_4 - \text{logsin } V_6 - \text{logsin } V_8 = 0\end{aligned} \tag{112}$$

Equations (111) and (112) are original condition equations which state the requirement which must be satisfied by the adjusted values of the angles, and correspond to the form shown in (56). It remains to derive the simpler *reduced* condition which expresses the relation between the *corrections* to the observed angles, so that it may be combined with the angle equations (105) into one adjustment having the same unknowns, namely, the v's. That is, for each V must be substituted its value, $M+v$, and the equation reduced to the linear form to correspond to (59), page 58.

If a given angle, M, be altered by a small correction, v, expressed in seconds, the logarithmic sine of the angle would be changed by a corresponding amount, namely, the number of seconds in v multiplied into the difference for one second in that particular logarithmic sine, as taken from the logarithmic tables with the proper algebraic sign, positive if the angle lie in the first quadrant, in which the sine increases with increasing angle, and negative if in the second quadrant, where the sine decreases with increase of angle. That is,

$$\text{logsin } (M+v) = \text{logsin } V = \text{logsin } M + v\ (d1'') \qquad (113)$$

For example, if $M = 76°\ 15'\ 14.5''$ and $v = -4.1''$, logsin $M =$ 9.9873797 with a difference for $1''$ of $+5.1$ in the seventh decimal place. Then logsin $(M+v) =$ logsin $76°\ 15'\ 10.4'' = 9.9873797 - 4.1(+5.1) = 9.9873776$. Substituting for logsin V, in (112), its value given in (113), namely, logsin $M + (d\,1'')v$, and collecting the logarithms into one constant term, we have,

$$\begin{aligned}(d_1 1'')v_1 - (d_2 1'')v_2 + (d_3 1'')v_3 - (d_4 1'')v_4 + (d_5 1'')v_5 \\ - (d_6 1'')v_6 + (d_7 1'')v_7 - (d_8 1'')v_8 + (\text{logsin } M_1 \\ - \text{logsin } M_2 + \text{logsin } M_3 - \text{logsin } M_4 + \text{logsin } M_5 \\ - \text{logsin } M_6 + \text{logsin } M_7 - \text{logsin } M_8) = 0 \qquad (114)\end{aligned}$$

in which $(d_1 1'')$ represents the difference for $1''$ in the logsine of angle M_1, in the seventh place of decimals, assuming that seven-place logarithms are used.[1] Each of these differences for $1''$ is a numerical coefficient for its v, and corresponds to a, b, c, etc., of (59), page 58. Also, since the observations are carefully made, the observed angles, M, will approximate closely to their adjusted values, V, so that the algebraic sum of the logsines of the M's in (114) will be, instead of zero as in (112), a small error of closure, q, expressed in units of the seventh decimal place, and equal to the amount by which the sum of the positive logsines exceeds that of the negative ones, in (114). The *reduced form* of the side equation becomes, therefore, if we assume it to be the fourth of the condition equations so that its coefficients are d's (a, b, and c being coefficients of the first three conditions respectively),

$$d_1v_1 - d_2v_2 + d_3v_3 - d_4v_4 + d_5v_5 - d_6v_6 + d_7v_7 - d_8v_8 + q_4 = 0 \qquad (115)$$

in which each d is the numerical difference for $1''$ in the logsine of the corresponding angle. Thus we see that the side equation

[1] It will be convenient, generally, to use seven-place logarithms but to take the sixth place unit for $(d1'')$ and q, thus moving the decimal point one place to the left in these quantities. If the tables show a difference for $10''$ to be $+51$, therefore, $(d1'') = +5.1$ in the seventh place, or $+0.51$ in the sixth place.

states that the corrections to the angles must be such that the algebraic sum of the resulting corrections to their logsines will equal $-q$ and the original side equation (111) be satisfied. Since the coefficients and the constant term of (115) are expressed in the same unit, in the seventh place, this equation is consistent with the angle equations (105), as stating a linear relation between the v's.

94. A Shorter Form of the Side Equation for the quadrilateral may be obtained by computing one of the *adjacent* sides, as *BC*, from *AB*, instead of the opposite one, *CD*, as above. In Fig. 13,

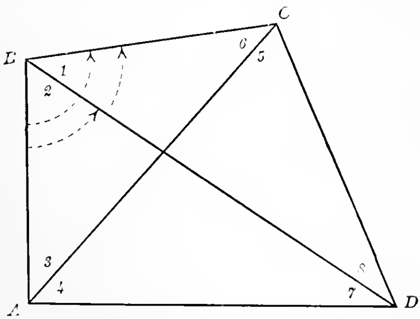


FIG 13. Side Equation; Quadrilateral.

the dotted arrows show the two routes of the computation, from which the two resulting values for *BC* must be equated:

$$BC = AB\frac{\sin V_3}{\sin V_6}$$

and

$$BC = AB\frac{\sin (V_3+V_4)}{\sin V_7}\cdot\frac{\sin V_8}{\sin (V_5+V_6)} \qquad (116)$$

whence,

$$\frac{\sin V_3 \sin (V_5+V_6) \sin V_7}{\sin (V_3+V_4) \sin V_6 \sin V_8} = 1 \qquad (117)$$

and the reduced side equation becomes,

$$d_3v_3 - d_{3+4}(v_3+v_4) + d_{5+6}(v_5+v_6) - d_6v_6 + d_7v_7 - d_8v_8 + q_4 = 0 \quad (118)$$

or, separating the various unknowns,

$$(d_3 - d_{3+4})v_3 - d_{3+4}v_4 + d_{5+6}v_5 + (d_{5+6} - d_6)v_6$$
$$+ d_7v_7 - d_8v_8 + q_4 = 0 \qquad (119)$$

in which d_{3+4} represents the difference for 1″ in the logsine of the sum-angle, M_3+M_4, etc. Although this form is somewhat shorter than (115) in that the angles at the station, B, lying between the sides AB and BC, do not appear, it is more troublesome for the beginner. However, the fact that the angles at one station are not concerned makes this the preferable form when one of the stations of the figure was not completely occupied, in which case the equation is expressed between the two exterior lines adjacent to this station. Thus, in the above figure, station B might have been unoccupied without affecting the form of equation (117).

95. Side Equation for a Central-point Figure. Let Fig. 14

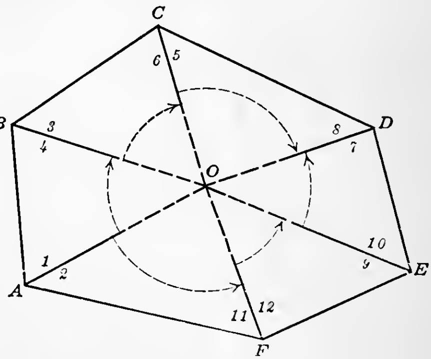


FIG. 14. Side Equation; Central-point Figure.

represent the general form of a central-point figure, and for the sake of variety, suppose the central station to have been observed from each of the others but not to have been occupied, as shown by the lines being broken at that point, but that otherwise the figure is complete. The side equation will be written between two of the lines which meet at the central point, and the dotted arrows show the two routes of computation from the line AO to DO through two series of triangles.

$$DO = AO\frac{\sin V_1 \sin V_3 \sin V_5}{\sin V_4 \sin V_6 \sin V_8} = AO\frac{\sin V_2 \sin V_{12} \sin V_{10}}{\sin V_{11} \sin V_9 \sin V_7} \tag{120}$$

whence,

$$\frac{\sin V_1 \sin V_3 \sin V_5 \sin V_7 \sin V_9 \sin V_{11}}{\sin V_2 \sin V_4 \sin V_6 \sin V_8 \sin V_{10} \sin V_{12}} = 1 \tag{121}$$

The reduced side equation follows directly, as in Art. 93:

$$d_1v_1 - d_2v_2 + d_3v_3 - d_4v_4 + d_5v_5 - d_6v_6 + d_7v_7$$
$$- d_8v_8 + d_9v_9 - d_{10}v_{10} + d_{11}v_{11} - d_{12}v_{12} + q = 0 \qquad (122)$$

It will be seen that the odd-numbered angles occur in the numerator of (121) and the even ones in the denominator, which results, as in Art. 93, from the *clockwise numbering* of the two angles at each occupied station. Since the angles at the central station were not observed, and do not occur in the side equation, it is not necessary to number them. By comparison, also, with (111) of Art. 93, it is evident that the side equation for a central-point figure having four sides and the stations *A*, *B*, *C*, *D*, and *O*, would be identical with (111) written for a complete quadrilateral with diagonals.

96. Mechanical Statement of Side Equations. The similarity among the side equations (111), (117), and (121), in the occurrence of the odd-numbered angles in the numerators and the even ones in the denominators, would indicate the possibility of writing these equations by inspection instead of using this property merely as a check. This may be done by the following mechanical method for all ordinary figures of the quadrilateral or central-point form.

Notation. (*a*) The *pole* for a side equation is some station, or other point, to which a line runs from every (other) station of the figure for which the equation is to be written. It may be at the intersection of the diagonals of a quadrilateral, although there is no station there. The point selected as the pole will be indicated by a *small circle* drawn around it, as in the following figures:

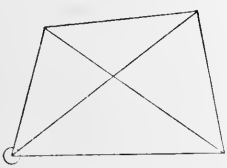


FIG. 15.

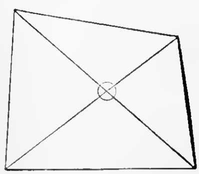


FIG. 16.

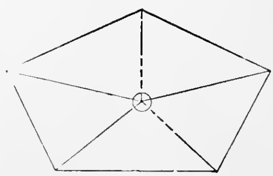


FIG. 17

Location of Pole for Side Equation.

(*b*) At each station of the figure, there will be three lines, one of which goes to the pole and may be called the *pole line*. Of the other two, one will be the *left-hand line* and the other, the *right-hand line,* as we look into the figure, from the station towards the pole. (*c*) At each station, the *left-hand angle* is the angle between the left-hand line and the pole line, and the *right-hand angle* is the one between the right-hand line and the pole line. This notation is illustrated in the following diagrams in which l and r indicate the left-hand and right-hand lines and L and R, the left-hand and right-hand angles, respectively:

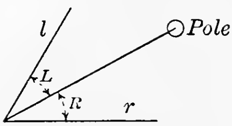


Fig. 18.

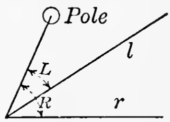


Fig. 19.

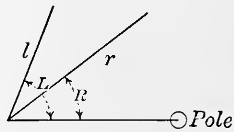


Fig. 20.

Left-hand and Right-hand Angles.

The *side equation,* then, is written by placing the product of the sines of the left-hand angles equal to that of the sines of the right-hand angles, or by placing the former in the numerator and the latter in the denominator of a fraction which is placed equal to unity. The reduced form of the equation, corresponding to (115), (119), and (122), may be written as the sum of the dv's for the left-hand angles minus the dv's for the right-hand angles plus q equals zero, or,

$$[dv](\text{for left-hand angles}) - [dv](\text{for right-hand angles}) + q = 0 \quad (123)$$

in which d is the difference for $1''$ in the logsine and q is the sum of the logsines of the left-hand angles minus the sum of the logsines of the right-hand angles.

It will be noted that the angles at the pole do not enter into the side equation at all, so that the pole is situated at the intersection of the two lines between which the equation would be written according to the analytical method of the preceding articles. Thus, equation (111) would correspond to a pole at the intersection of the diagonals of Fig. 12, page 88; in Fig. 13, page 91, the pole would be at station B for equation (117); and in Fig. 14, page 92, the pole would be at the central station, O.

It is now apparent, as stated in Arts. 93 and 95, that the odd-numbered angles occurred on one side of the side equations and the even ones on the other because the two angles at each exterior station of the figures had been numbered clockwise so that the odd ones were on the left and the even ones on the right of the pole lines.

The selection of the pole may be governed by the following principles.[1] In a central-point figure, it must be at the central station. If a station was not occupied, or not completely occupied, the pole should be at that station. Sum-angles may be avoided by placing the pole of a quadrilateral at the intersection of the diagonals, which is simpler for the beginner although it introduces two additional angles and logsines. The pole should not be placed where the smallest angles occur, as these angles should enter into the side equations with their larger coefficients.[2]

97. Number of Side Equations in a Figure. In order that there may be two routes, through two series of triangles, between two lines in a figure, there must be at least three triangles in the figure, and therefore, four stations with three lines to each station. In other words, the quadrilateral is the simplest figure for which a side equation may be written. Similarly, a central-point figure, without diagonals, can have but one side equation since there are but two series of triangles through which one side may be computed from another. The quadrilateral and the central-point figure, therefore, furnish one side equation each, and are the elemental figures for these equations.

A complete central-point figure has as many outer lines and the same number of inner ones as there are exterior stations, so that the total number of lines will equal twice the total number of stations less one, which is true, also, in a quadrilateral. Since each of these figures yields one side equation, the formula may be written,

$$\text{Number of Side Equations} = L - 2(S-1) + 1 = L - 2S + 3 \qquad (124)$$

in which L is the total number of lines, full or broken, and S is the

[1] See Wright and Hayford's Adjustment of Observations.

[2] The logsine of a small angle varies rapidly with change of angle, so that the difference for 1″ is large.

total number of stations whether occupied or not. Adding to either figure one station with three lines from it to stations of the figure, adds another quadrilateral or central-point figure, and therefore, another side equation, which corresponds to an increase of one $(L-2S)$ in the formula. Each additional line, without increase of stations, makes possible the writing of one or more new side equations, using that line, of which, however, only one can be regarded as *independent*. Adding one station to a figure thus adds as many side equations as there are lines from that station *to the figure*, less two. For example, adding to Fig. 21 the station

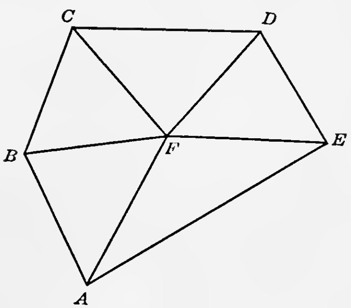


Fig. 21.

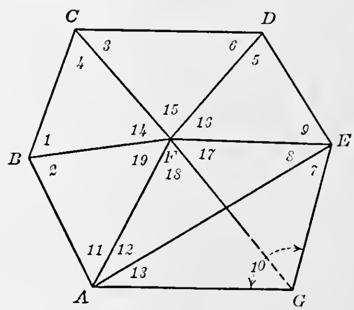


Fig. 22.

Number of Side Equations in a Figure.

G with three lines to A, F, and E, adds one side equation which could be written for the quadrilateral G–A–F–E or for the central-point figure G–A–B–C–D–E, and if this latter had been the original figure, the addition of the line AE would have formed the quadrilateral A–F–E–G with its side equation. In each case the new side equation must include the added station or line.

98. Statement of All of the Conditions for a Figure Adjustment. In the preceding articles, the three kinds of conditions, local, angle, and side, which enter into the adjustment, have been explained separately. The adjustment as a whole will now be considered.

Strictly, all of the conditions should be satisfied simultaneously in one general adjustment. The labor of computation is greatly reduced, however, by diminishing the number of conditions, and it is especially convenient to perform the local adjustments separately and in advance of the figure adjustment, inasmuch as it is good practice to arrange the observations so as to have but one local condition at a station, involving all of the angles, as explained in Art. 87. Therefore, we shall assume that the necessary local adjustments have been made, as in Arts. 79 and 82, preparatory to the figure adjustment. However, if the angles at any station of the figure complete the horizon, it will be necessary to include in the figure adjustment a local condition providing that the sum of these angles shall remain 360°, that is, that the algebraic sum of the corrections to these angles must be zero. This is likely to be the case at an interior station, such as F, in Fig. 22. Also, if a sum-angle should be included among the conditions, as well as its component angles, and with a separate number, a similar local condition would be necessary to insure that the sum-angle would remain equal to the sum of its components after adjustment; but this may well be avoided by designating the sum-angle as the sum of its components, as in Art. 94, instead of using a separate symbol for it.[1] In general, care must be taken that the preliminary adjustment be not disturbed by the later one.

The selection of the angle and side equations for a given figure or system must conform to the requirements that all the necessary conditions be included, but no more, and that they be independent of one another, so that no one of them could be obtained by combining any of the others. If a dependent condition were included, by mistake, it would be indicated during the solution of the normal equations by a derived equation having all of its coefficients zero, or nearly so, so that the corresponding correlate would be indeterminate. The necessary number of independent angle and side equations will be given by formulas (107) and (124), namely,

$$\text{Number of Angle Equations} = L' - S' + 1 \tag{107}$$

$$\text{Number of Side Equations} = L - 2S + 3 \tag{124}$$

in which L and S are the total numbers of lines and stations, and

[1] These local conditions are avoided in the figure adjustment by using *directions* instead of *angles*, as will be shown later on.

L' is the number of full lines and S' is the number of occupied stations. (For a station to be considered as occupied, at least two lines must be unbroken at that station.) The best method of writing the angle and side equations so as to be certain of their independence as well as their number, is to draw a sketch of the system or figure to be adjusted, adding one station at a time, with its lines to the previous stations, and writing the equations introduced by that station and those lines. For each station so added, there will be as many angle equations as new *full* lines, less one, and as many side equations as new lines, less two. As has been stated, small angles should be used in the side equations where practicable, although it is best to use each but once. In angle equations, on the contrary, they should be avoided.

For example, the equations for Fig. 22, page 96, will be written. In this case, $L=13$, $L'=12$, $S=S'=7$, and there are six angle and two side equations. The complete horizon at F, moreover, requires a local condition. Beginning with the line AB, station F, with the two lines to A and B, forms a triangle with one angle equation (A), as shown below. Adding C with two lines to B and F, gives one angle equation, (B), and similarly, adding D with two lines to C and F gives (C). Now, with E are added three lines to A, D, and F, so that two angle equations, (D) and (E), are formed and one side equation (H), for the whole figure A–B–C–D–E–F, with pole at F. With G are added two full lines and one broken line, giving one angle equation, (F), for the triangle G–A–E, and one side equation (I) which might well be written for the quadrilateral G–A–F–E, with pole at G since the line FG is broken at G. Thus we have six angle and two side equations as required by the formulas above. The local equation (G) for the station F must be added. To facilitate the formation and solution of the normal equations, these condition equations are so arranged as to place the simpler ones first and the more complex ones with larger coefficients, last. (See Art. 78, page 71.) The angle equations, therefore, will usually precede the side equations. For a central point figure, also, several angle equations may be written in succession having no angles in common, with the result that many of the coefficients in the first normal equations will be zero, thus materially reducing the labor of solution. The above conditions are arranged as follows:

$$
\begin{array}{lll}
\text{Angle:} & (A) & V_2 + V_{11} + V_{19} - (180° + \epsilon_a) = 0 \\
& (B) & V_1 + V_4 + V_{14} - (180° + \epsilon_b) = 0 \\
& (C) & V_3 + V_6 + V_{15} - (180° + \epsilon_c) = 0 \\
& (D) & V_5 + V_9 + V_{16} - (180° + \epsilon_d) = 0 \\
& (E) & V_8 + V_{12} + V_{17} - (180° + \epsilon_e) = 0 \\
& (F) & V_7 + V_{10} + V_{13} - (180° + \epsilon_f) = 0 \\
\text{Local:} & (G) & V_{14} + V_{15} + V_{16} + V_{17} + V_{18} + V_{19} - 360° = 0 \\
\text{Side:} & (H) & \dfrac{\sin V_1}{\sin V_2} \dfrac{\sin V_3}{\sin V_4} \dfrac{\sin V_5}{\sin V_6} \dfrac{\sin V_8}{\sin V_9} \dfrac{\sin V_{11}}{\sin V_{12}} = 1 \\
& (I) & \dfrac{\sin V_7 \sin V_{17} \sin (V_{12} + V_{13})}{\sin (V_7 + V_8) \sin V_{13} \sin V_{18}} = 1
\end{array}
\qquad (125)
$$

Substituting for each V its $M+v$, the M's being the observed values of the angles, and computing the spherical excesses, the angle and local equations are thrown into their reduced form as in (105), and the reduced side equations are formed as in (115) and (119), respectively. The formation of the nine normal equations and the remainder of the solution then follow as in the last chapter.

99. Adjustment of a Quadrilateral: Method of Angles. To illustrate the foregoing principles, the following quadrilateral

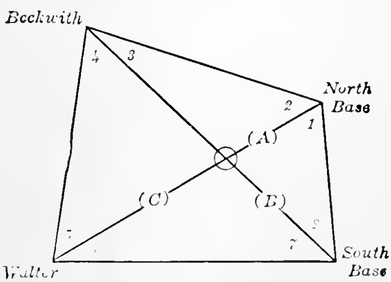


FIG. 23. Adjustment by Method of Angles.

will now be adjusted in full. The final angles are desired to hundredths of a second. Weights are equal. Seven-place logarithms will be used but with the unit taken in the sixth place for the side equation. The pole is taken at the intersection of the diagonals, as indicated. The given angles are shown in the three triangles which will be used in succession for the angle equations.

GIVEN ANGLES

(A)

Beckwith..........	(3)	26°	42′	51.8″
North Base........	(1)	64	43	42.3
North Base........	(2)	43	44	02.0
South Base........	(8)	44	49	27.4
				03.5

$\epsilon_a = 0.05''$ $q_a = +3.45''$

(B)

Walter............	(6)	28°	17′	12.9″
North Base........	(1)	64	43	42.3
South Base........	(7)	42	09	40.3
South Base........	(8)	44	49	27.4
				02.9

$\epsilon_b = 0.06''$ $q_b = +2.84''$

(C)

Walter............	(5)	48°	03′	10.3
Walter............	(6)	28	17	12.9
Beckwith..........	(4)	61	29	53.9
South Base........	(7)	42	09	40.3
				57.4

$\epsilon_c = 0.08''$ $q_c = -2.68''$

SIDE EQUATION

	(M)			Logsin (+)	d1″		(M)			Logsin (−)	d1″
1	64°	43′	42.3″	9.9563098	+ 9.9	2	43°	44′	02.0″	9.8396728	+22.0
3	26	42	51.8	9.6527716	+41.9	4	61	29	53.9	9.9438915	+11.4
5	48	03	10.3	9.8714339	+19.0	6	28	17	12.9	9.6756749	+39.1
7	42	09	40.3	9.8268640	+23.3	8	44	49	27.4	9.8481489	+21.2
				9.3073793						9.3073881	
										9.3073793	
										—88	

(Unit is in seventh place; to be changed to sixth.)

CONDITION EQUATIONS

	v_1	v_2	v_3	v_4	v_5	v_6	v_7	v_8	(q)
(a)	+1	+1	+1					+1	+3.45
(b)	+1					+1	+1	+1	+2.84
(c)				+1	+1	+1	+1		−2.68
(d)	+0.99	−2.20	+4.19	−1.14	+1.90	−3.91	+2.33	−2.12	−8.80
(s)	+2.99	−1.20	+5.19	−0.14	+2.90	−1.91	+4.33	−0.12	

NORMAL EQUATIONS

	A	B	C	D	(q)	Check	Sum
(1)	+4	+2	0	+ 0.86	+3.45	+10.31	+ 6.86
(2)		+4	+2	− 2.71	+2.84	+ 8.13	+ 5.29
(3)			+4	− 0.82	−2.68	+ 2.50	+ 5.18
(4)				+53.50	−8.80	+42.03	+50.83
(2)		+4	+2	− 2.71	+2.84	+ 8.13	(Factors)
		−1	0	− 0.43	−1.72	− 5.16	−2/4 = −0.50
(II)		+3	+2	− 3.14	+1.12	+ 2.9~~7~~8	
(3)			+4	− 0.82	−2.68	+ 2.50	
			0	0	0	0	0
			−1.33	+ 2.09	−0.75	− 1.99	−2/3 = −0.667
(III)			+2.67	+ 1.27	−3.43	+ 0.51 ✓	
(4)				+53.50	−8.80	+42.03	
				− 0.18	−0.74	− 2.22	−0.86/4 = −0.215
				− 3.29	+1.17	+ 3.12	+3.14/3 = +1.047
				− 0.60	+1.63	− 0.24	−1.27/2.67 = −0.476
(IV)				+49.43	−6.74	+42.69 ✓	

CORRELATES

$$D = \frac{+6.74}{49.43} = +0.136$$

$$C = \frac{+3.43 - 0.17}{2.67} = \frac{+3.26}{2.67} = +1.221$$

$$B = \frac{-1.12 + 0.43 - 2.44}{3} = \frac{-3.13}{3} = -1.043$$

$$A = \frac{-3.45 - 0.12 + 0 + 2.09}{4} = \frac{-1.48}{4} = -0.370$$

TESTS OF CORRELATES

	Eq. (4)	Eq. (3)	Eq. (2)	Eq. (1)
A	−0.32	0.00	−0.74	−1.48
B	+2.83	−2.09	−4.17	−2.09
C	−1.00	+4.88	+2.44	0.00
D	+7.28	−0.11	−0.37	+0.12
(q)	−8.80	−2.68	+2.84	+3.45
Sum	—0.01	0.00	0.00	0.00

CORRECTIONS

	v_1	v_2	v_3	v_4	v_5	v_6	v_7	v_8
aA	−0.370	−0.370	−0.370					−0.370
bB	−1.043					−1.043	−1.043	−1.043
cC				+1.221	+1.221	+1.221	+1.221	
dD	+0.135	−0.299	+0.570	−0.155	+0.258	−0.532	+0.317	−0.288
Sum	−1.278	−0.669	+0.200	+1.066	+1.479	−0.354	+0.495	−1.701
Corr'n	−1.28	−0.67	+0.20	+1.06	+1.48	−0.35	+0.49	−1.70

TESTS OF CORRECTIONS

(A) $-1.28-0.67+0.20-1.70+3.45=0.00$
(B) $-1.28-0.35+0.49-1.70+2.84=0.00$
(C) $+1.06+1.48-0.35+0.49-2.68=0.00$
(D) $-1.27+1.47+0.84-1.21+2.81+1.37+1.14+3.60-8.80=-0.05$

(This discrepancy is 5 in the eighth place, and negligible.)

TEST OF SIDE EQUATION

	(V)				(V)	
1	64° 43′ 41.02″	9.9563085		2	43° 44′ 01.33″	9.8396713
3	26 42 52.00	9.6527724		4	61 29 54.96	9.9438927
5	48 03 11.78	9.8714367		6	28 17 12.55	9.6756735
7	42 09 40.79	9.8268652		8	44 49 25.70	9.8481454
		9.3073828				9.3073829

Computation of Triangles

The ultimate test of the adjustment occurs in the computation of the lengths of the lines or sides of the triangles. If an error were made in the original side equation, such as an erroneous logsine or difference for 1″, or an error in adding the angles of a triangle to obtain its error of closure, q, all of the subsequent operations might check, to and including the tests of corrections. The test of the side equation, using the adjusted angles and the new logsines, checks the original logsines and differences for 1″. It remains to be seen in the computation of triangles whether or not the adjusted angles "fill" each of the four triangles and at the same time satisfy the side equation by giving the same results for lengths which are computed in two triangles. The above discrepancy of one in the last place of logarithms in the side equation test, would show, also, in the triangle computations, but is too small to warrant further investigation. It would be corrected arbitrarily so as to leave no inconsistency in the computed results. (An example of the final triangle computations will be given at the close of the adjustment of this quadrilateral by the Method of Directions, which follows.)

100. Use of Directions instead of Angles. In the measurement of angles with a direction instrument, as in primary triangulation, the various signals are sighted independently and for each pointing the horizontal circle is read, in a clockwise direction. This is done in the direct and reversed positions of the instrument and in various positions of the circle, and the mean of all of the readings upon a certain signal is adopted as the *direction* to that signal. The angle between any two signals is the direction of the right-hand signal minus that of the left-hand one, and there is no local adjustment. Even though the separate angles be measured by reading directions in pairs, or by the method of repetitions, the directions may be numbered, instead of the angles, and each angle designated by the difference of the two directions which limit it, the right-hand one minus the left. In Fig. 24, for example, angle BAC would be designated by the symbols $-(1)+(2)$, and CAE would be represented by $-(2)+(4)$. EAB would be $-(4)+(1)$, and so always minus the left plus the right-hand direction, clockwise.

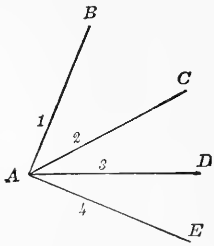


FIG. 24. Directions.

This method has certain advantages, especially in the adjustment of the more complex systems, which render its use very desirable, and it is deservedly popular among computers. One of its strongest features lies in the fact that preliminary local adjustments are not disturbed by later adjustments in which the method of directions is used, so that no local condition for an interior station would be necessary in a case such as that in Art. 97 and Fig. 22. Each direction is regarded as observed independently, and the unknowns of the problem are the corrections to the separate directions. The correction to an angle, therefore, would be the correction to the right-hand direction minus that of the left one, algebraically. There will be more directions, in a given system, than angles, but this is not a serious objection when the Method of Correlates is used. (In the Method of Indirect Observations, any increase in the number of unknowns produces a like increase in the normal equations but in the Method of Correlates the number of normal equations is equal to that of the conditions.)

The weights of the directions will be equal, in the general case, but different weights may be assigned if certain signals were more difficult to observe than others, owing, perhaps, to unsteady atmospheric conditions or poor illumination. If it be desired to use directions in the adjustment of *angles of different weights*, care should be taken in giving weights to the corresponding directions that the weights of adjacent angles be not seriously affected. If two adjacent directions were assigned small weights, and thereby received large corrections, the intervening angle might receive a small correction (the difference of the two large ones) and so defeat the purpose of the computer. If two adjacent angles have small weight, the intervening direction might be given a smaller weight and thus affect both angles. However, if the angles have different weights, rather than the directions, it may be best to adjust by the Method of Angles explained above. In using directions, therefore, we shall assume that angle weights are equal; if separate directions have different weights, they may be treated exactly as in the Method of Angles.

If directions be used in local adjustments, it is advisable to use the Method of Indirect Observations, as in Art. 82, since the local conditions would be identities of the form,

$$-(1)+(2)-(2)+(3)-(3)+(4)-(4)+(1)-360°=0 \quad (126)$$

Therefore, it will usually be preferable to use the Methods of Angles and Correlates, illustrated in Art. 79, for the local adjustment.

101. Notation: Method of Directions. In numbering the directions of a figure, one side may be regarded as the initial line, as if fixed by a previous adjustment, perhaps, and its numbers omitted. In this case, it is well to place letters on the fixed line, instead of numbers, to distinguish its directions, when writing the equations, these lettered directions not to enter into the reduced conditions, and to receive no corrections. On the other hand, this use of letters is not necessary, and numbers may be placed upon all of the directions, if desired, without altering the method or increasing the work to any considerable extent. Directions are to be numbered clockwise, invariably, at each station, so as to avoid errors. Unobserved directions, shown by broken lines, will not be numbered.

102. Lists of Directions. Preparatory to the adjustment, a list of the directions at each station may be made from the given data. The names of the observed signals are arranged *in clockwise order*, usually beginning with a prominent one which is given the initial direction, 0° 00′ 00.0″, although any direction may be taken as the zero line. If angles were observed, and adjusted locally, if necessary, the resulting angles are added in the proper order so as to obtain the angle from the assumed initial direction to each of the other signals or stations, which will be its direction in the list. The angle from one station to another, clockwise, will then be the direction of the latter minus that of the former. If only two or three angles were observed at each station, it may not be worth while to form these lists of directions, but each angle may be given the proper designation as the difference between two directions, and two angles added or subtracted, when necessary, to obtain a third.

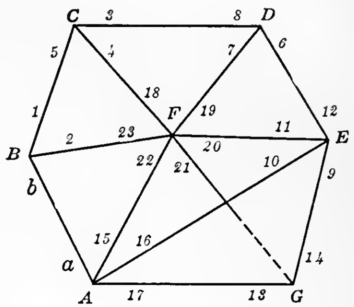


FIG. 25. Adjustment by Method of Directions.

103. Statement of Conditions: Method of Directions. To illustrate the use of directions, the condition equations (125) for Fig. 22, page 96, will be restated. The figure is reproduced in Fig. 25, with the directions numbered. The line AB will be regarded as fixed and its directions will be indicated by the letters a and b, merely for convenient reference. Comparison of the following equations with (125) will make the process evident. As stated above, the local condition is unnecessary when the Method of Directions is used, as the closure of the horizon at F will not be disturbed by applying corrections to the directions, since each direction is common to two angles and its correction must increase one by the same amount as it decreases the other, leaving the sum unchanged.

$$
\begin{array}{llllllll}
(A) & -V_2 & +(b) & -(a) & +V_{15} & -V_{22} & +V_{23} & -(180^\circ+\epsilon_a)=0 \\
(B) & -V_1 & +V_2 & -V_4 & +V_5 & -V_{23} & +V_{18} & -(180^\circ+\epsilon_b)=0 \\
(C) & -V_3 & +V_4 & -V_7 & +V_8 & -V_{18} & +V_{19} & -(180^\circ+\epsilon_c)=0 \\
(D) & -V_6 & +V_7 & -V_{11} & +V_{12} & -V_{19} & +V_{20} & -(180^\circ+\epsilon_d)=0 \\
(E) & -V_{10} & +V_{11} & -V_{15} & +V_{16} & -V_{20} & +V_{22} & -(180^\circ+\epsilon_e)=0 \\
(F) & -V_9 & +V_{10} & -V_{13} & +V_{14} & -V_{16} & +V_{17} & -(180^\circ+\epsilon_f)=0
\end{array}
\qquad (127)
$$

$$(G) \quad \frac{\sin(-V_1+V_2)\sin(-V_3+V_4)\sin(-V_6+V_7)\sin(-V_{10}+V_{11})\sin(-a+V_{15})}{\sin(-V_2+b)\sin(-V_4+V_5)\sin(-V_7+V_8)\sin(-V_{11}+V_{12})\sin(-V_{15}+V_{16})}=1$$

$$(H) \quad \frac{\sin(-V_9+V_{10})\sin(-V_{20}+V_{21})\sin(-V_{15}+V_{17})}{\sin(-V_9+V_{11})\sin(-V_{21}+V_{22})\sin(-V_{16}+V_{17})}=1$$

To state the reduced conditions, the direction letters, a and b, are omitted and the angle equations are written by replacing each V by its v, and $-(180°+\epsilon)$ by the closure error, q. The side equations are more complicated owing to the combined subscripts. For example,

$$\text{logsin}\,(-a+V_{15})=\text{logsin}\,(+V_{15})=\text{logsin}\,(+M_{15})+d_{+15}v_{15} \quad (128)$$

and

$$\begin{aligned}\text{logsin}\,(-V_3+V_4)&-\text{logsin}\,(-V_4+V_5)\\ &=\text{logsin}\,(-M_3+M_4)-\text{logsin}\,(-M_4+M_5)\\ &\qquad +d_{-3+4}(-v_3+v_4)-d_{-4+5}(-v_4+v_5)\\ &=\text{logsin}\,(-M_3+M_4)-\text{logsin}\,(-M_4+M_5)-d_{-3+4}v_3\\ &\qquad +(d_{-3+4}+d_{-4+5})v_4-d_{-4+5}v_5\end{aligned}$$

in which d_{-3+4} is the difference for $1''$ in the logsine of the angle $(-M_3+M_4)$, etc. Applying these principles to the equations (127), we obtain the reduced equations in the following form:

$$\begin{aligned}
&(A)\quad -v_2+v_{15}-v_{22}+v_{23}+q_a=0\\
&(B)\quad -v_1+v_2-v_4+v_5-v_{23}+v_{18}+q_b=0\\
&(C)\quad -v_3+v_4-v_7+v_8-v_{18}+v_{19}+q_c=0\\
&(D)\quad -v_6+v_7-v_{11}+v_{12}-v_{19}+v_{20}+q_d=0\\
&(E)\quad -v_{10}+v_{11}-v_{15}+v_{16}-v_{20}+v_{22}+q_e=0\\
&(F)\quad -v_9+v_{10}-v_{13}+v_{14}-v_{16}+v_{17}+q_f=0 \qquad (129)\\
&(G)\quad -d_{-1+2}v_1+(d_{-1+2}+d_{-2})v_2-d_{-3+4}v_3+(d_{-3+4}+d_{-4+5})v_4\\
&\qquad -d_{-4+5}v_5-d_{-6+7}v_6+(d_{-6+7}+d_{-7+8})v_7\\
&\qquad -d_{-7+8}v_8-d_{-10+11}v_{10}+(d_{-10+11}+d_{-11+12})v_{11}\\
&\qquad -d_{-11+12}v_{12}+(d_{+15}+d_{-15+16})v_{15}-d_{-15+16}v_{16}+q_g=0\\
&(H)\quad (-d_{-9+10}+d_{-9+11})v_9+d_{-9+10}v_{10}-d_{-9+11}v_{11}\\
&\qquad -d_{-20+21}v_{20}+(d_{-20+21}+d_{-21+22})v_{21}-d_{-21+22}v_{22}\\
&\qquad -d_{-15+17}v_{15}+(d_{-15+17}-d_{-16+17})v_{17}+d_{-16+17}v_{16}+q_h=0
\end{aligned}$$

Inspection of equations (127) shows that there are more unknowns than in (125), and what is more important, that adjacent angle equations have unknowns in common so that less of the product-coefficients in the normal equations would be zero in this case than in the other. The diagonal coefficients (squares) are larger, also, owing to the greater number of unknowns. These are disadvantages which may offset the omission of the local condition, in a central-point figure, so that the Method of Angles might actually involve less work than the Method of Directions, in such a case. A rearrangement of the above equations, however, would simplify the normal equations, to some extent, by collecting the zero coefficients near the beginning. The following order might be used: (*B*), (*D*), (*F*), (*E*), (*C*), (*A*), (*H*), (*G*).

104. Adjustment of a Quadrilateral: Method of Directions.

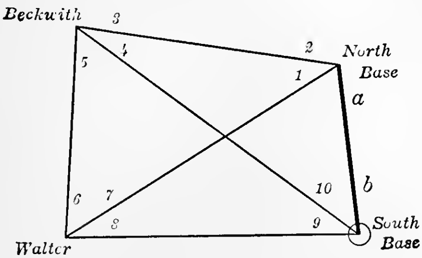


FIG. 26. Adjustment of Quadrilateral; Method of Directions.

As an example of the use of directions, the quadrilateral of Art. 99 will be adjusted. A comparison of the two methods of solving the same problem will be instructive. The figure is shown in Fig. 26 with the new notation. The number of angles being small, it will not be necessary to write lists of directions, but the separate angles of the triangles will be represented by the proper directions, instead, and other angles may be obtained from them by addition or subtraction, the symbols being subjected to the same operations. Thus, adding the two angles, $-(3)+(4)$ and $-(4)+(5)$, we obtain their sum as $-(3)+(5)$. The pole for the side equation is taken at South Base. The right-hand angles happen to have been written on the left side of the equation, and vice versa, which is equivalent to changing all the signs in the equation without affecting the results. The computation of the triangles is added in order to complete the solution.

(*A*)

Beckwith.	$-(3)+(4)$	26°	42′	51.8″
N. Base.	$-(a)+(2)$	108	27	44.3
S. Base.	$-(10)+(b)$	44	49	27.4
				03.5

$\epsilon_a = 0.05''$ $\qquad q_a = +3.45''$

(*B*)

Walter.	$-(7)+(8)$	28°	17′	12.9″
N. Base	$-(a)+(1)$	64	43	42.3
S. Base.	$-(9)+(b)$	86	59	07.7
				02.9

$\epsilon_b = 0.06''$ $\qquad q_b = +2.84''$

(*C*)

Walter.	$-(6)+(8)$	76°	20′	23.2″
Beckwith.	$-(4)+(5)$	61	29	53.9
S. Base.	$-(9)+(10)$	42	09	40.3
				57.4

$\epsilon_c = 0.08''$ $\qquad q_c = -2.68''$

SIDE EQUATION

(Right-hand Angles)		Logsine e (+)	$d1''$	(Left-hand Angles)		Logsine e (−)	$d1''$
−(4)+(5)	61° 29′ 53.9″	9.9438915	+11.5	−(3)+(4)	26° 42′ 51.8″	9.6527715	+41.8
−(a)+(2)	108 27 44.3	9.9770521	− 7.0	−(a)+(1)	64 43 42.8	9.9563099	+ 9.9
−(7)+(8)	28 17 12.9	9.6756750	+39.1	−(6)+(8)	76 20 23.2	9.9875381	+ 5.1
						−7.5966195	
		+7.5966186				+7.5966186	
						$q_d =$ −9	

(Unit is in the seventh place; it will be changed to the sixth.)

CONDITION EQUATIONS

	v_1	v_2	v_3	v_4	v_5	v_6	v_7	v_8	v_9	v_{10}	(q)
(a)		+1	−1	+1						−1	+3.45
(b)	+1						−1	+1	−1		+2.84
(c)				−1	+1	−1		+1	−1	+1	−2.68
(d)	−0.99	−0.70	+4.18	−5.33	+1.15	+0.51	−3.91	+3.40			−0.90
(s)	+0.01	+0.30	+3.18	−5.33	+2.15	−0.49	−4.91	+5.40	−2.00	0	

NORMAL EQUATIONS

	A	B	C	D	(q)	Check	Sum
(1)	+4	0	−2	−10.21	+3.45	− 4.76	− 8.21
(2)		+4	+2	+ 6.32	+2.84	+15.16	+12.32
(3)			+6	+ 9.37	−2.68	+12.69	+15.37
(4)				+75.78	−0.90	+80.36	+81.26
(2)		+4	+2	+ 6.32	+2.84	+15.16	(Factors)
(5)		0	0	0	0	0	0
(II)		+4	+2	+ 6.32	+2.84	+15.16	
(3)			+6	+ 9.37	−2.68	+12.69	
(6)			−1	− 5.10	+1.72	− 2.38	+2/4
(7)			−1	− 3.16	−1.42	− 7.58	−2/4
(III)			+1	+1.11	−2.38	+ 2.73√	
(4)				+75.78	−0.90	+80.36	
(8)				−26.06	+8.81	−12.15	+10.21/4
(9)				− 9.99	−4.49	−23.95	− 6.32/4
(10)				− 0.30	+0.66	− 0.75	− 1.11/4
(IV)				+39.43	+4.08	+43.51√	

CORRELATES

$$D=\frac{-4.08}{+39.43}=-0.103$$

$$C=\frac{+0.114+2.375}{4}=+0.622$$

$$B=\frac{-1.244+0.654-2.84}{4}=-0.858$$

$$A=\frac{0+1.244-1.056-3.45}{4}=-0.815$$

TESTS OF CORRELATES

	Eq. (1)	Eq. (2)	Eq. (3)	Eq. (4)
A	−3.26	0	+1.63	+8.32
B	0	−3.43	−1.72	−5.42
C	−1.24	+1.24	+3.73	+5.83
D	+1.06	−0.66	−0.96	−7.83
(q)	+3.45	+2.84	−2.68	−0.90
	+0.01	−0.01	0	0

Corrections

	v_1	v_2	v_3	v_4	v_5	v_6	v_7	v_8	v_9	v_{10}
aA		−0.815	+0.815	+0.815						+0.815
bB	−0.858						+0.858	−0.858	+0.858	
cC				−0.622	+0.622	−0.622		+0.622	−0.622	+0.622
dD	+0.102	+0.072	−0.433	+0.552	−0.119	−0.053	+0.405	−0.352		
Sum	−0.756	−0.743	+0.382	−0.885	+0.503	−0.675	+1.263	−0.588	+0.236	+1.437
(v)	-0.76^5	−0.74	+0.38	-0.88^9	+0.50	−0.68	+1.26	−0.59	+0.24	+1.44

Tests of Corrections

(A)	−0.74	−0.38	-0.88^9	−1.44	+3.45	$= +0.01^0$				
(B)	-0.76^5	−1.86	−0.59	−0.24	+2.84	$= -0.01^0$				
(C)	$+0.88^9$	+0.50	+0.68	−0.59	−0.24	+1.44	−2.68	$= -0.01^0$		
(D)	+0.75	+0.52	+1.59	$+4.69^{74}$	+0.58	−0.35	−4.92	−2.00	−0.90	$= -0.0\cancel{4}+0.01$

Test of Side Equation

$-(4)+(5)$	61°	29′	55.29″	9.9438931	$-(3)+(4)$	26°	42′	50.53″	9.6527661
$-(a)+(2)$	108	27	43.56	9.9770527	$-(a)+(1)$	64	43	41.55	9.9563091
$-(7)+(8)$	28	17	11.05	9.6756675	$-(6)+(8)$	76	20	23.29	9.9875381
				7.5966133					7.5966133 ✓

COMPUTATION OF TRIANGLES

	Stations	Obs'd Angles (M)	Corr'ns (v)	(V)	Excess (ϵ)	Plane Angles and Distances	Logs
	N. Base—S. Base					3503.140 m.	3.5444574
3+1	Beckwith	26° 42′ 51.8″	−1.27″	50.53″	0.02″	26° 42′ 50.51″	0.3472338
a+2	N. Base	108 27 44.3	−0.74	43.56	0.01	108 27 43.55	9.9770527
−10+b	S. Base	44 49 27.4	−1.41	25.96	0.02	44 49 25.94	9.8481459
	Beckwith—N. Base				0.05	5493.348+1	3.7398371+1
	S. Base—Beckwith					7391.693	3.8687439
	N. Base—S. Base					3503.140	3.5444574
−7+8	Walter	28° 17′ 12.9″	−1.85″	11.05″	0.02″	28° 17′ 11.03″	0.3243324
−a+1	N. Base	64 43 42.3	−0.75	41.55	0.02	64 43 41.53	9.9563091
−9+b	S. Base	86 59 07.7	−0.24	07.46	0.02	86 59 07.44	9.9993986
	Walter—N. Base				0.06	7382.244	3.8681884
	S. Base—Walter					6684.962	3.8250989
	S. Base—Beckwith					7391.693	3.8687439
−6+8	Walter	76° 20′ 23.2″	+0.09″	23.29″	0.02″	76° 20′ 23.27″	0.0124619
−1+5	Beckwith	61 29 53.9	+1.39	55.29	0.03	61 29 55.26	9.9438931
−9+10	S. Base	42 09 40.3	+1.20	41.50	0.03	42 09 41.47	9.8268667
	Walter—Beckwith				0.08	5105.902+2	3.7080725+1
	S. Base—Walter					6684.962	3.8250989
	N. Base—Walter					7382.244	3.8681884
3+5	Beckwith	88° 12′ 45.7″	+0.42″	45.82″	0.02″	88° 12′ 45.80″	0.0002114
1+2	N. Base	43 44 02.0	+0.01	02.01	0.03	43 44 01.98	9.8396728
−6+7	Walter	48 03 10.3	+1.94	12.24	0.02	48 03 12.22	9.8714375
	Beckwith—N. Base				0.07	5493.351−2	3.7398373−1
	Walter—Beckwith					5105.904	3.7080726

In this standard form of computation of the triangle sides, the given side is written first, followed by the opposite station and the other two in clockwise order around the triangle. The corrections are applied to the given angles to obtain the adopted (spherical) ones, from which the spherical excesses are deducted and the plane angles (to be used in the logarithmic computation) are found. The sum of these plane angles, of course, should be exactly 180°. The cologsine of the first angle is written below the log distance, followed by the logsines of the other two angles. Covering the fourth logarithm with a pencil or strip of paper, the first three are added to obtain the sixth, and the fifth is then obtained as the sum of the first, second, and fourth, by covering the third. In order that the computed lengths may be consistent throughout, a certain value is adopted for each distance and logarithm, and the necessary modifications are made by the application of small corrections as shown. It is a good plan to arrange the triangles in the above form before beginning the adjustment of the figure. Then the symbols and the observed angles (after local adjustment, if any) are in convenient form for use, together with the spherical excesses. After the adjustment, the corrections are inserted and the form completed.

105. Adjustment of a Quadrilateral: Approximate Method.[1] The angles of a quadrilateral may be made to satisfy the angle equations exactly and the side equation very nearly, by an approximate adjustment which, although not rigorous, may be sufficient for subordinate triangulation or detached figures in which great precision is not required. The weights of the angles are assumed to be equal.

The two triangles formed by one diagonal are first closed by correcting the four angles of each by one-fourth of the closure-error for the triangle. One of the other triangles is then closed by correcting each of its four new angles by one-fourth of its closure-error, which correction is also applied to the remaining four angles (of the fourth triangle), with the opposite sign, so that all four triangles are thus satisfied exactly. Taking the pole for the side

[1] Due to Prof. T. W. Wright.

equation at the intersection of the diagonals, each of the eight new angles is corrected by one-eighth of the error of closure of the log-sines divided by the algebraic mean of the eight differences for $1''$, the angles on the right being corrected with the opposite sign to those on the left, so as to bring the sums of their logsines closer together. If the eight angles were equal, the side equation, also, would be exactly satisfied by this method; in this case the figure would be a square.

For example, let us adjust the quadrilateral in Fig. 23, page 99, with the data and notation there given. (See next page.)

106. Adjustment to Conform to Work Previously Adjusted or Fixed. Triangulation of a subordinate character is frequently carried on in connection with a main scheme or net in order that a number of points may be located from the main stations without reoccupying them expressly for this purpose. In primary triangulation, for instance, it is customary to read directions from the stations upon prominent objects such as church-spires, which may be used later by local surveyors for obtaining initial positions and azimuths. Such points do not enter into the adjustment of the main figures but are adjusted subsequently and usually separately, upon the previously adjusted work as a basis. Also, secondary or tertiary figures may be connected to or based upon primary ones so as to require separate adjustment which will not disturb the previous work. If the connection be to one fixed line, only, that line would be used as a base-line, and no condition would be introduced. But if a triangle be fixed, or two sides and the included angle, the new conditions must be so written as not to disturb the previous adjustment. The Method of Directions is particularly convenient when fixed lines are involved, as the directions may be omitted from those lines and they will not be affected by the adjustment. The angles are assumed to have been adjusted locally, in advance. The use of letters on the fixed lines, instead of numbers, serves to identify them without giving them the character of unknown directions, although an experienced computer usually omits the letters as well as the numbers on the fixed lines. If the Method of Angles were used, local conditions

ADJUSTMENT OF A QUADRILATERAL; APPROXIMATE METHOD

	Obs'd Angles				(+)	(−)	d 1″	(Final)
(4)	61° 29′ 53.9″	54.57″		55.18″		9.9438930	+11.4	54.81″
(5)	48 03 10.3	10.97		11.58^{2}	9.8714363		+19.0	11.96
(6)	28 17 12.9		13.57	12.96		9.6756752	+39.1	12.59
(7)	42 09 40.3		40.97	40.36^{5}	9.8268641		+23.3	40.72
	179 59 57.4			00.08√				00.08√
	$\epsilon_c = 0.08$							
	2.68							
	+0.67							
(8)	44 49 27.4		26.54	25.93		9.8481459	+21.2	25.56″
(1)	64 43 42.3		41.44^{3}	40.82	9.9563084		+ 9.9	41.19
(2)	43 44 02.0	01.14		01.75		9.8396722	+22.0	01.38
(3)	26 42 51.8	50.94		51.55	9.6527705		+41.9	51.92
	180 00 03.5		02.51	00.05√	9.3073793	9.3073863	187.8	00.05√
	$\epsilon_a = 0.05$		$\epsilon_b = 0.06$			3793	+23.5	
	3.45		2.45			70	= Mean	
	−0.86		−0.61		70 ÷ 8 = 8.75 = average log corrn.			
Remainder	(−0.01)		(−0.01)		8.75 ÷ 23.5 = 0.37″ = corrn. to each angle			

TEST OF SIDE EQUATION

(1) 64° 43′ 41.19″	9.9563087	(2) 43° 44′ 01.38″	9.8396714
(3) 26 42 51.92	9.6527721	(4) 61 29 54.81	9.9438926
(5) 48 03 11.96	9.8714370	(6) 28 17 12.59	9.6756737
(7) 42 09 40.72	9.8268650	(8) 44 49 25.56	9.8481451
	9.3073828		9.3073828√

would have to be added. The following simple cases will be considered. From the condition equations the solution proceeds in the usual manner.

107. Two Sides and the Included Angle Fixed. Fig. 27. The adjacent sides, A and B, are fixed in length and the angle between them, also, must not be altered. If the missing diagonal had been

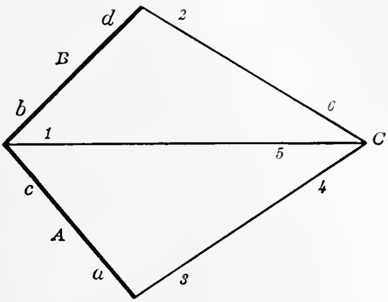


Fig. 27. Two Sides and Included Angle Fixed

observed, it would have to be considered as fixed, since the sides, A and B, and the included angle, determine the triangle completely. That case will be discussed later. The only new lines, then, are the three which run to C, forming the two triangles. According to our rules, there are two angle equations, one for each triangle, and no side equation. However, the fact that two lines are fixed in length renders a condition necessary, which shall require the angles to be so adjusted that when one fixed line is computed from the other, the result will be equal to its known length. This condition is called a *length equation*. It has the same nature as a side equation, but involves two known lengths. The three equations are:

$$(A) \qquad -(b)+V_1-V_2+(d)-V_5+V_6-(180°+\epsilon_a)=0$$

$$(B) \qquad -V_1+(c)-(a)+V_3-V_4+V_5-(180°+\epsilon_b)=0 \qquad (130)$$

$$(C) \qquad A\,\frac{\sin\,(-a+V_3)\,\sin\,(-V_5+V_6)}{\sin\,(-V_4+V_5)\,\sin\,(-V_2+d)}=B$$

It is obvious that the angle $(-a+V_3)$, for example, may be designated by $(+V_3)$ since there is to be no correction to the direction (a). The equations (130) may therefore be written,

$$\begin{aligned} &(A) \quad +V_1-V_2-V_5+V_6-(180°+\epsilon_a)=0 \\ &(B) \quad -V_1+V_3-V_4+V_5-(180°+\epsilon_b)=0 \\ &(C) \quad \frac{A \sin(+V_3)\sin(-V_5+V_6)}{B \sin(-V_4+V_5)\sin(-V_2)}=1. \end{aligned} \tag{131}$$

To obtain the error of closure, q, for the length equation, the logarithm of the length A must be added and that of B subtracted, in the series of logsines. As these lengths are fixed, they do not appear in the reduced conditions, which have the form,

$$\begin{aligned} &(A) \quad +v_1-v_2-v_5+v_6+q_a=0 \\ &(B) \quad -v_1+v_3-v_4+v_5+q_b=0 \\ &(C) \quad d_{-2}v_2+d_{+3}v_3+d_{-4+5}v_4-(d_{-4+5}+d_{-5+6})v_5+d_{-5+6}v_6+q_c=0 \end{aligned} \tag{132}$$

108. Quadrilateral with One Fixed Triangle. Fig. 28. The quadrilateral being complete would have one side and three angle equations. The angle equation for the fixed triangle is satisfied in advance, however, so that there remain but two angle equations, and the side equation as independent conditions. Using the tri-

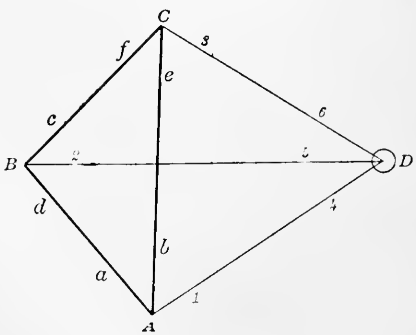


FIG. 28. One Triangle Fixed

angles $D-A-B$ and $D-B-C$, and placing the pole at D for the side equation, we write the three conditions:

$$\begin{aligned}
&(A) \quad -(a)+V_1-V_2+(d)-V_4+V_5-(180^\circ+\epsilon_a)=0 \\
&(B) \quad -(c)+V_2-V_3+(f)-V_5+V_6-(180^\circ+\epsilon_b)=0 \qquad (133) \\
&(C) \quad \frac{\sin(-a+V_1)\sin(-c+V_2)\sin(-V_3+e)}{\sin(-b+V_1)\sin(-V_2+d)\sin(-V_3+f)}=0
\end{aligned}$$

After obtaining the constants, q, the lettered directions, a, b, c, d, e, and f, would be omitted, as usual, in forming the reduced equations, although it is convenient to use them in the subscripts of the side equation to distinguish between those angles which differ only by the fixed angle at A, B, or C. Thus the reduced side equation would be,

$$\begin{aligned}
(C) \quad (d_{-a+1}-d_{-b+1})v_1+(d_{-c+2}+d_{-2+d})v_2 \\
-(d_{-3+e}-d_{-3+f})v_3+q_c=0 \quad (134)
\end{aligned}$$

109. Fixed Triangle or Polygon with Central Point Unoccupied. Figs. 29 and 30. In this case there are no new triangles and, therefore, no angle equations whatever. The pole for the single side equation is placed at the central, unoccupied (or concluded) station. If this station be outside of the triangle, the side equation would be the same as (C) in (133) and (134), above, but if it be inside the figure, the side equation has a characteristic symmetry

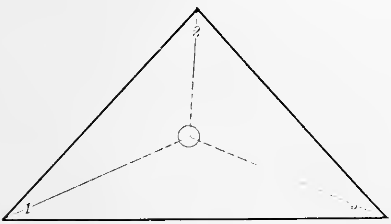


FIG. 29.

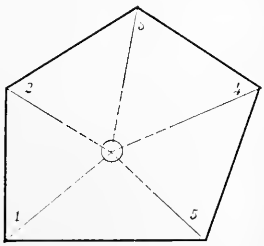


FIG. 30.

Fixed Triangle or Polygon with Concluded Station

in that every numbered direction occurs in both numerator and denominator, and the algebraic signs are positive in the numerator and negative in the denominator, or vice versa. Thus, for Fig. 29, the side equation would be as follows, omitting the lettered directions which are unnecessary in this very simple case,

$$\frac{\sin(+V_1)\sin(+V_2)\sin(+V_3)}{\sin(-V_1)\sin(-V_2)\sin(-V_3)}=1 \qquad (135)$$

Similarly, for Fig. 30, the side equation would be,

$$\frac{\sin(+V_1)\sin(+V_2)\sin(+V_3)\sin(+V_4)\sin(+V_5)}{\sin(-V_1)\sin(-V_2)\sin(-V_3)\sin(-V_4)\sin(-V_5)}=1 \qquad (136)$$

This case, especially Fig. 29, occurs so frequently in the location of subordinate stations that it may well receive special attention here. In its reduced form, equation (135) may be written,

$$(d_{+1}+d_{-1})v_1+(d_{+2}+d_{-2})v_2+(d_{+3}+d_{-3})v_3+q=0 \qquad (137)$$

in which d_{+1} is the difference for 1″ in the logsine of angle $(+M_1)$, d_{-1} is that difference for angle $(-M_1)$, etc. This equation has the same form as (86) of Art. 80, page 73, so that $a_1=(d_{+1}+d_{-1})$, $a_2=(d_{+2}+d_{-2})$, etc. Assuming equal weights, which will usually be the case, the correlate for the single equation will be, from (87),

$$A = \frac{-q}{[aa]} \qquad (138)$$

and the corrections will follow from (88),

$$v_1 = a_1A = a_1\frac{-q}{[aa]}; \quad v_2 = a_2A = a_2\frac{-q}{[aa]}; \text{ etc.} \qquad (139)$$

It is easy, then, to arrange the logsines in positive and negative columns, and to take their algebraic sum as q. The algebraic sum of the differences for 1″ corresponding to the directions (those in the negative column having their signs changed) will be the a's, and the sum of the squares of these a's is the denominator of the factor, $-q/[aa]$, in (139). Each v is computed by multiplying its a into this factor.

For example, let us adjust the following observed angles for Fig. 29, the weights being equal. Each angle is followed by its logsine and the difference for 1″ in that logsine, the left-hand angles, in the left-hand column, being considered positive. The spherical excess for the fixed triangle is 0.30″.

(+)

$+M_1$	27°	39′	42.1″	9.6667519	+40.1
$+M_2$	30	32	58.3	9.7061055	+35.6
$+M_3$	31	34	10.9	9.7189460	+34.3
	89	46	51.3	9.0918034	
	90	13	09.0		
	180	00	00.3	(Sum of angles of fixed triangle.)	

(−)

$-M_1$	31°	49′	17.6″	9.7220376	+33.9
$-M_2$	25	43	07.2	9.6374423	+43.8
$-M_3$	32	40	44.2	9.7323383	+32.8
	90	13	09.0	9.0918182	
				+9.0918034	
				$q=-148$	

The unit will be taken in the sixth place instead of the seventh in order to diminish the coefficients.

$$
\begin{array}{lr}
 & a^2 \\
a_1=(d_{+1}+d_{-1})=+4.01+3.39=+7.40 & 54.76 \\
a_2=(d_{+2}+d_{-2})=+3.56+4.38=+7.94 & 63.04 \\
a_3=(d_{+3}+d_{-3})=+3.43+3.28=+6.71 & 45.02 \\
\hline
 & [aa]=162.82
\end{array}
$$

The reduced side equation is,

$$+7.40v_1+7.94v_2+6.71v_3-14.8=0$$

Thus, $-q\ [aa]=+14.8/162.82=+0.091''$, and the corrections are,

$$v_1=+7.40(+0.091'')=+0.67''$$

$$v_2=+7.94(+0.091'')=+0.72''$$

$$v_3=+6.71(+0.091'')=+0.61''$$

The final, adjusted angles and their logsines (to test the work) are,

$+V_1$	27° 39′ 42.77″	9.6667546	$-V_1$	31° 49′ 16.93″	9.7220353
$+V_2$	30 32 59.02	9.7061081	$-V_2$	25 43 06.48	9.6374392
$+V_3$	31 34 11.51	9.7189481	$-V_3$	32 40 43.59	9.7323363
	89 46 53.30	9.0918108		90 13 07.00	9.0918108 ✓
	90 13 07.00				
	180 00 00.30 ✓				

The side equation is satisfied, since the sums of the positive and negative logsines are equal, namely, 9.0918108. Also, the sum of all the angles remains unchanged, and each of the three angles of the fixed triangle is the same as before the adjustment, since each correction was applied both positively and negatively. An ordinary slide-rule is sufficient for the arithmetical work, and after the sum of the *aa*'s is obtained, each *v* is found at one setting of the rule. The above illustration of the process is given in greater detail than is necessary when the method is understood.

110. Adjustment of a System between Points of Control. Large systems of triangulation, such as the primary work of the U. S. Coast and Geodetic Survey, may extend over strips of country for hundreds of miles. In such great distances, errors of various kinds are likely to have a cumulative effect which becomes too great to be tolerated. It is necessary, therefore, to control, or check, the triangulation at intervals which will depend upon the precision of the observations, the points of control being farthest apart in first-class, or primary systems. The *lengths* are controlled by measured base-lines, the *positions*, by astronomical observations for azimuth, latitude, and longitude, and the *elevations*, by precise spirit leveling, although the astronomical observations may really control all three elements of a system, that is, its size, shape, and position. (See Art. 85, page 80.)

In general, the controlling points for these different purposes will not be coincident. The observations for azimuth may not be made at the same stations as those for latitude or longitude, or at the base-line stations. To illustrate the character of the control, however, it will be assumed for example that a given system, or net, starts at a certain line, *AB*, Fig. 31, whose length and azimuth are known as well as the latitude, longitude, and elevation of one of its ends, and that it extends to another line, *CD*, which is fixed in the same manner, in length, direction, position, and elevation. This line, *CD*, may have been fixed by original observations and measurements, as if it were a detached or isolated line, or it may be a line in a previously adjusted triangulation system which is so precise or so strong that it is not subject to modification by the subsequent work.

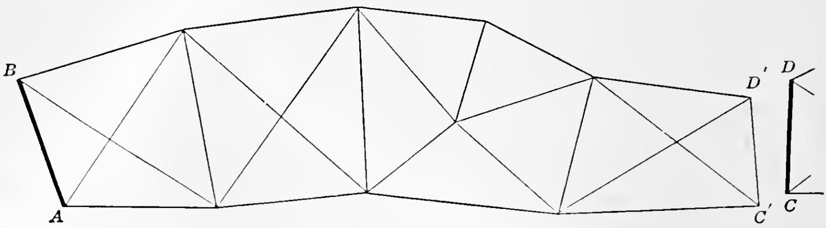


FIG. 31. Triangulation System with Control

If the separate elemental figures, such as quadrilaterals, are adjusted in advance, with local, side, and angle equations, and the lengths and positions are computed from the initial side, AB, through the system, the final line might fall at $C'D'$ instead of CD. If, then, all the lines of the system were flexible except AB, and $C'D'$ were picked up and forced to coincide with CD, it is easily seen that all of the lines and angles would probably be distorted. The adjustment, therefore, affects all of the angles in the net. The *Base-line*, or *Length, Equation* provides that the length of $C'D'$, computed from AB, shall be equal to the fixed length, CD. This condition is similar to the length equation (C) of Art. 107, page 121, but must extend through the whole net. The *Azimuth Equation* requires that $C'D'$ shall be parallel to CD. The *Latitude Equation* states that the latitude of a point such as C' shall be equal to the fixed latitude of the corresponding point, C, and the *Longitude Equation* expresses the same requirement for their longitudes. It is evident that these conditions are independent—that any one or more of them could be satisfied without forcing the others to be fulfilled. For example, the line $C'D'$ might have the same length as CD, and C' might coincide with C, and still the azimuths might be different. Of course, the amount of the discrepancy is exaggerated in the figure.

In a rigid adjustment, all of these conditions would be combined with the local, side, and angle conditions and satisfied simultaneously. It is usually sufficient, and much more convenient, however, to perform the figure adjustments and then modify them so as to effect the closure upon the controlling line through the above four conditions. In primary triangulation of the highest grade, the rigid, complete adjustment may be required.

When an extensive system contains several points of control, such as base-lines or astronomical stations, it is customary to regard the net as subdivided at these points into sections, and to adjust each section independently. This method has practical advantages which outweigh its divergence from the ideal adjustment of an entire system as a single problem.[1]

The special case sometimes occurs in which a detached net, having its own base-line, but only approximate astronomical position, is connected, after its figure adjustment, to a fixed system through a single figure, such as a quadrilateral or central-point figure. If the discrepancy in length between the two systems be small, it may be thrown entirely into the intervening figure, which would have, therefore, two lines fixed in length, as shown in Figs. 32 and 33. In addition to the usual angle and side equations, the figure would have a length equation such as (C) of Art. 107, page 121. It is assumed that the geographic positions for the detached system are to be obtained, through this connection, from the fixed one.[2]

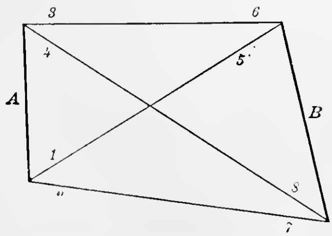


Fig. 32. Length Equation

[1] For a thorough treatment of the adjustment of large systems, and for special methods applicable to triangulation in general, see Wright and Hayford's Adjustment of Observations; Special Publication No. 28 of the U. S. Coast and Geodetic Survey, by O. S. Adams; and Jordan's Vermessungskunde, Band I.

[2] A method for the adjustment of triangulation by correcting the preliminary latitudes and longitudes of the stations is presented by Mr. Adams in Special Publication No. 28 of the U. S. Coast and Geodetic Survey.

111. Adjustment of Trigonometric Leveling. The adjustment of the horizontal angles in triangulation is generally independent of the vertical angles, which will be used to compute the difference of elevation between the various stations. Although

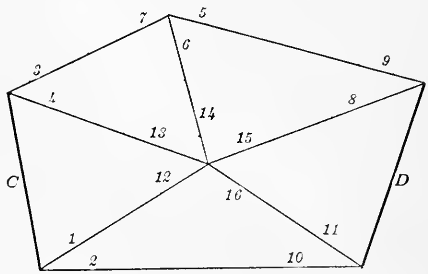


FIG. 33. Length Equation

these vertical angles may be adjusted directly, it is usually easier and at the same time satisfactory, to adjust the computed differences of elevation, and this is done by the method illustrated in Art. 77, page 64. A long net may be divided into sections to facilitate the adjustment, and if a control point becomes available in the form of a station whose elevation has been determined directly through a line of precise levels, the entire net intervening between the initial elevation and this final one may be adjusted to conform to this total difference of elevation by a slight modification of the partial adjustments without disturbing their conditions, as the proportionate discrepancies will be very small in carefully executed work.

112. Base-lines. The measurement of a base-line is carried out in sections, and the total length is the sum of the sections. It is customary to measure each section two or more times, in both directions and under different conditions. The separate measures of a section are then adjusted as direct observations, by taking their mean or weighted mean. It is seldom that weights are required, however, since additional measures are made if there is too much discrepancy between the first two, and doubtful results are subject to rejection in the field.

CHAPTER VII

EMPIRICAL FORMULAS

113. Empirical Formulas. Experimental investigations frequently comprise the determination of the values of a certain function corresponding to known, assigned, or observed values of its independent variable. It is often desirable to express the relation thus determined, between the function and the variable, in the form of an equation. Should the observations be the same in number as the unknown constants or coefficients of the equation, a rigid solution of the problem would result, as explained in Art. 22. But, as it is customary to make a larger number of observations in order to obtain increased precision, the problem becomes one of determining the equation which will best represent the entire group of observations, thus involving an adjustment by the Method of Least Squares. Such an expression, depending upon experimental data, is known as an *Empirical Formula.*

114. Their Uses. Empirical formulas are sometimes called *interpolation formulas* from the fact that one of their principal uses is to facilitate the interpolation of values of the function among the observations. The curve which represents the formula is smooth and continuous and avoids the discrepancies among the various observations, so that interpolation is usually safe and reasonable. However, there is generally a tendency to use the formula beyond the limits of the observations, that is, to extrapolate along an extension of the curve. While this yields, in many cases, very useful and interesting results, care must be taken that such results be not considered trustworthy except within reasonable limits.

Sometimes it seems impossible to derive a theoretical relation between two variables, while it is evident from the observations that some connection does exist. Here the empirical formula may be the only one available.

It is not essential that the relation expressed by the function have any foundation in theory. It may be purely accidental, as is the case in many statistical investigations. A formula may be stated between the death-rate of a city and the time or season, or between the depth of a pond and the distance from the shore. However, the existence of a close relationship, such as cause and effect, is sometimes indicated by an empirical formula, resulting in the subsequent development of the rigid formula by theoretical analysis. In this manner some well-known laws have been discovered.

115. Nature of the Problem. Equations may be partially or wholly empirical. For example, the *form* of the expression may be developed theoretically and regarded as known, leaving only the numerical constants and coefficients to be obtained empirically. Or, nothing whatever may be known concerning the formula, in which case it is necessary to assume a form for the equation and then determine the constants by an adjustment of the observations. In any event, the problem is, to ascertain those constants which will make the given expression, whether of previously known or assumed form, represent the observations as nearly as possible. Should there be uncertainty as between different forms which could be assumed, or should the residuals resulting from a solution be unsatisfactorily large, one or more other forms may be assumed and the constants be determined for each of them, that one being finally adopted for which the sum of the squares of the residuals is the least.

116. The Form of the Equation may be known from theoretical considerations, as when it is a special case of a group of expressions the nature of which is known. But in the great majority of cases, it must be obtained from the observations themselves. This is conveniently done by plotting them as rectangular coordinates, representing the values of the function, y, as ordinates, and those of the independent variable, x, as abscissas, each point thus plotted corresponding to one observation. A smooth curve is then sketched so as to follow the plotted points as nearly as practicable. An inspection of this curve will generally throw it into one of three classes, namely: (1) a portion of a conic section, such as a straight line or a parabola; (2) a periodic or wave-like curve; or (3) a curve which is non-linear with respect to the unknown coefficients, that is, one which involves their products, squares, or higher powers, or functions.[1]

To assist in the selection of a suitable form, a number of curves, with their equations, are shown in Appendix E. Apparent properties of the desired curve should be carefully noted, as positions of axes, asymptotes, points of inflection, points of crossing of axes, maxima and minima, regular or irregular periodicity, etc., so that the equation selected may be capable of representing these features. In general, however, it is convenient to utilize an expression in the form of a series which can embrace all the curves in a certain group. This is particularly useful in the first two of the above classes of curves, and will now be illustrated.

117. Straight Lines and Parabolic Arcs. The simpler curves vary from the straight line, through the forms which appear uniformly curved, to those in which the sharpness of curvature increases or decreases continuously in one direction. It is possible to represent any of these by a series of the form,

$$y=a+bx+cx^2+dx^3+ex^4+ \ . \ . \ . \tag{140}$$

The character of the curve will determine the number of terms to be used in this equation. Thus, if a straight line be desired, the first two terms would suffice, giving,

$$y=a+bx \tag{141}$$

If the curvature is slight, or if the curve straightens towards one end, the parabolic form may be assumed,

$$y=a+bx+cx^2 \tag{142}$$

Or, if it be desired to represent the plotted points still more closely, one or more terms may be added, the principle being that the greater the number of terms used, the more nearly will the resulting formula fit the observations. If an unnecessarily large number of terms is used, the coefficients of those which might be omitted

[1] It must be remembered that in the derivation of empirical formulas, the variables, x and y, are not the unknowns as they are in the Adjustment of Indirect Observations, Chapter III. Here, the variables are the observed quantities and the coefficients are the unknowns which are to be determined. As will appear later, the methods of solution are analogous.

will come out quite small or negligible, and a re-solution with the simpler form may be advisable.

It should be noted that the straight line is a special case, and that although the plotted points seem to lie very close to such a line it is usually best to use the formula for a parabola and obtain a curve which approximates closely to the straight line. This parabolic form is of very general application for empirical formulas because of its convenience and adaptability.

118. Periodic Functions. If the curve is composed of similar elements which repeat themselves as x increases, the function is evidently periodic, that is, the values of y corresponding to increasing values of x will pass through similar cycles or periods. The curve in many cases will have a wave-like form, and it may be simple or very complex. The general formula to be used is a Fourier series,

$$y=a+b\sin\frac{360^\circ}{m}x+c\cos\frac{360^\circ}{m}x$$
$$+d\sin\frac{360^\circ}{m}2x+e\cos\frac{360^\circ}{m}2x+\ .\ .\ . \qquad (143)$$

in which a, b, c, etc., are the constants to be determined.[1] By using a sufficient number of terms, this equation may represent any curve whatever, for finite values of the variables, but in the case of periodic functions it is particularly useful. If the elementary parts of the curve are alike and not complicated, the first three terms will be sufficient; otherwise, succeeding pairs of terms should be added, involving the multiples of x. Unless a complex formula is expected, it is well to sketch each wave in the curve so it will be symmetrical about its middle ordinate. If the total angle corresponding to a cycle should be 180° instead of 360°, this number should be substituted for the latter in the formula.

The constant quantity, m, is the number of units of x in one cycle or period, and is assumed from an inspection of the curve and the observations. For example, suppose the brightness of a

[1] For an interesting application of harmonic analysis to this problem, see Brunt's Combination of Observations, Chapter XI

variable star to be observed from day to day, and when plotted as a function of the time to seem to have a period of about nine days. Here, x would be the number of days elapsed since an assumed epoch (such as the date of the first observation) and m would be assumed as 9. Thus, x/m is an abstract number; $360°x/m$ is a number of degrees; and $360°/m$ is a constant coefficient of x in any single problem. Different values of m may be assumed, and the problem solved for each, if deemed worth while, that one being adopted for which the sum of the squares of the residuals is the least. In determining the period, m, from the curve, it is well to measure it at several places, if possible, and take the mean.

When the empirical formula of this periodic type has been determined, it may be transformed into a more convenient expression in the following manner: Let n_0, n_1, n_2, X_1, X_2, etc., be auxiliary quantities determined from the assumptions,

$$n_0 = a; \qquad n_1 \sin X_1 = b; \qquad n_2 \sin X_2 = d;$$
$$n_1 \cos X_1 = c; \qquad n_2 \cos X_2 = e; \text{ etc.} \tag{144}$$

Substituting in (143), and combining, we have

$$y = n_0 + n_1 \cos\left(\frac{360°}{m}x - X_1\right) + n_2 \cos\left(\frac{360°}{m}2x - X_2\right) + \ . \ . \ . \tag{145}$$

which is shorter than (143). From (144),

$$n_1 = \frac{b}{\sin X_1}, \qquad \tan X_1 = \frac{b}{c}, \text{ etc.}$$

119. Non-linear Forms. As stated in Art. 38, equations of higher degree can be reduced to linear form, in general, by Taylor's Theorem, and in the case of exponential equations by the use of logarithms. Thus, it is not necessary to treat these higher degree expressions except by reducing to the linear form and then applying the usual methods. These processes of reduction will now be explained. They are applicable, of course, to equations which are non-linear as to the independent variable as well as to those which are non-linear as to the coefficients.

120. Exponential Functions. Equations in which the unknown constant occurs as an exponent constitute a special case for reduction to linear form which, owing to its simplicity, will be discussed first. In brief, the method is to throw the equation into the logarithmic form, by taking the logarithm of each member, and the resulting function will be linear with respect to the desired coefficient. Suppose the function to be of the form

$$y=ax^b, \tag{145}$$

in which a and b are to be determined so as to fit all of the observations as well as possible. Taking the logarithm of each member,

$$\log y=\log a+b \log x \tag{146}$$

which has the linear form

$$y'=A+bx' \tag{147}$$

where A and b are the unknown constants.

By plotting log x and log y as coordinates, or by using logarithmic cross-section paper for plotting x and y, the above exponential formula would be represented by a straight line. Thus the assumption of this form of equation can be easily checked.

Special attention must be given to the *weights* in this case of exponential functions, for the weights of the reduced, linear equations will not be the same as before reduction to the linear form, even though they were then equal.[1] If the weights of the original observations of y_1, y_2, y_3, etc., are w_1, w_2, w_3, etc., the corresponding weights of the functions, log y_1, log y_2, log y_3, etc., will be $y_1^2w_1$, $y_2^2w_2$, $y_3^2w_3$, etc.[2] Or, if the original weights are equal, the reduced equations will be weighted directly as the squares of the corresponding observed values of y. If the empirical formula

[1] This matter was first brought to the attention of the author several years ago, by Mr. C. E. Van Orstrand.

[2] It will be shown in the next chapter (Art. 143) that the weights are inversely as the squares of the mean square errors, and that (Art. 152) the mean square error of a function of y is equal to the mean square error of y multiplied by the derivative of the function with respect to y. Thus,

$$\epsilon_{\log y}=\epsilon_y\frac{d\,(\log y)}{dy}=\epsilon_y\frac{1}{y} \tag{147a}$$

and

$$w_{\log y}=w_y y^2 \tag{147b}$$

the mean square errors being represented by ϵ.

follows the observations very closely, however, as is usually the case, these weights will not have much effect. In fact, the errors of observation may warrant neglecting them in most cases.

121. General Case of Reduction to Linear Form. A simple example of an equation of the non-linear form with respect to the coefficients would be the following:

$$y=a^2+b^3x+cx^2+d^3x^3+\ .\ .\ . \tag{148}$$

Thus, each observation equation would be a function of a, b, c, d, etc., since x and y would be the observed numerical quantities, so that if the observed values of the function, y, be represented as usual by M_1, M_2, M_3, . . . M_n, the observation equations would have the form,

$$\begin{aligned} f_1(a, b, c, \ .\ .\ .\ ) &= M_1 \\ f_2(a, b, c, \ .\ .\ .\ ) &= M_2 \\ f_3(a, b, c, \ .\ .\ .\ ) &= M_3 \\ .\ .\ .\ .\ .\ .\ &.\ .\ .\ .\ . \\ f_n(a, b, c, \ .\ .\ .\ ) &= M_n \end{aligned} \tag{149}$$

The functions f_1, f_2, f_3, etc., on the left-hand side of these equations are different owing to their having different numerical values of x. Now let the best or most probable values of a, b, c, etc., namely, those which will result from this solution, be A, B, C, etc., and let A_0, B_0, C_0, etc., represent approximate values of A, B, C, etc., determined by the solution of some of the observation equations as simultaneous equations. Then let

$$\begin{aligned} A &= A_0+a' \\ B &= B_0+b' \\ C &= C_0+c' \\ .\ .\ &.\ .\ . \end{aligned} \tag{150}$$

in which a', b', c', . . . are small corrections to the assumed approximate values, to be determined by this solution. The observation equations may now be written,

$$\begin{aligned} f_1(A_0+a',\quad B_0+b',\quad C_0+c',\ .\ .\ .\ ) &= M_1+v_1 \\ f_2(A_0+a',\quad B_0+b',\quad C_0+c',\ .\ .\ .\ ) &= M_2+v_2 \\ f_3(A_0+a',\quad B_0+b',\quad C_0+c',\ .\ .\ .\ ) &= M_3+v_3 \\ \dots\dots\dots\dots\dots\dots\dots\dots \end{aligned} \tag{151}$$

the residuals being represented by v_1, v_2, v_3, . . .

These functions will now be expanded by Taylor's Theorem. The unknown corrections, a', b', c', . . . being small, it is permissible to neglect the terms involving their products and higher powers. The constant terms, $f_1(A_0, B_0, C_0, \ldots)$, $f_2(A_0, B_0, C_0, \ldots)$, etc., will be combined with the corresponding M_1, M_2, etc., and represented by l_1, l_2, etc., thus,

$$f_1(A_0,\ B_0,\ C_0,\ .\ .\ .\ )-M_1=l_1 \tag{152}$$

The observation equations will then become,

$$\begin{aligned} l_1+\frac{df_1}{dA_0}a'+\frac{df_1}{dB_0}b'+\frac{df_1}{dC_0}c'+\ .\ .\ . &= v_1 \\ l_2+\frac{df_2}{dA_0}a'+\frac{df_2}{dB_0}b'+\frac{df_2}{dC_0}c'+\ .\ .\ . &= v_2 \\ \dots\dots\dots\dots\dots\dots\dots\dots \end{aligned} \tag{153}$$

which are linear with regard to a', b', c', . . . The differential coefficients are obtained by differentiating the left-hand members of (149) with respect to a, b, c, etc., and then substituting for these quantities their approximate values, A_0, B_0, C_0, etc. If now we let the differential coefficients be represented by a_1, b_1, c_1, etc., with the subscripts of the corresponding equations, we obtain,

$$\begin{aligned} a_1a'+b_1b'+c_1c'+\ .\ .\ .\ +l_1 &= v_1 \\ a_2a'+b_2b'+c_2c'+\ .\ .\ .\ +l_2 &= v_2 \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ a_na'+b_nb'+c_nc'+\ .\ .\ .\ +l_n &= v_n \end{aligned} \tag{154}$$

which are similar to (18), page 27. Normal equations having been formed as in (21) or (22), their solution in the usual manner results in the desired corrections, a', b', c', etc., which applied to the approximate values, A_0, B_0, C_0, etc., as in (150), give the most probable values, A, B, C, etc. From these, the desired non-linear coefficients of the original equation are computed directly, giving finally the empirical formula sought.

If the observations are of different weight, the general form of normal equations, (21), would be used as in Indirect Observations, Chapter III.

122. Determination of the Constants. The plotted observations having been investigated and a suitable form selected for the equation, reduced, if necessary, to the linear form as just explained, it remains to form the observation equations and from them the normal equations, the solution of which is to give the desired constants for the empirical formula. In general, it is similar to the case of Indirect Observations, and the methods of Chapter III are applicable. The function will be stated in the explicit form, $y=f(x)$, although, of course, these quantities may be reversed, if desired, to fit the conditions, into $x=f(y)$, which form may sometimes be simpler than if fractional exponents were used.

The observation equations are formed, one for each observation, by substituting for x and y their observed values. The processes of Arts. 48 and 49 may be utilized for the simplification of the equations, and the normal equations will take the form of (22) or (21) according as the weights are equal or unequal. The solution of the normal equations will be carried out by the usual methods, and the resulting values of the unknowns, modified as necessary, will furnish the constant term and coefficients of the empirical formula.

123. Test of Empirical Formula. There are two methods of determining how closely the formula corresponds to the observations, namely, by plotting the curve of the formula and by computing the residuals.

The residuals are formed by substituting the observed values of the variable, x, in the empirical formula and computing the corresponding values of y. Subtracting from these the observed values of y, we obtain the residuals with the signs of corrections to the observations. The sum of the squares of these residuals is the quantity which should be a minimum if the empirical formula is the most probable one.

Having plotted the values of y, computed as above from the formula, the need of other, intermediate values in order to accurately define the curve may be seen at once and such values computed and plotted and the curve drawn by means of a French curve. If this be done on the sheet showing the original observations, the value of each residual is shown to scale by the *vertical* distance from the corresponding observation up or down to the curve, according as the residual is plus or minus, measured on its ordinate. Inspection of these graphical residuals will determine whether or not another form of curve should be assumed and the work repeated in order to find a closer approximation to the observations. If this should be done, the sums of the squares of the residuals in the two cases would be compared and that formula adopted for which this sum is the smaller. In a case of great importance, especially one that involves a large number of observations, several trials of this kind might be made in order to obtain the best formula.

124. Remarks. The above method of deriving empirical formulas is evidently closely analogous to the adjustment of Indirect Observations, that is, observations of a function of several quantities, and it must be borne in mind that in this method the errors of observation are assumed to lie in the values of the function, y, and not in those of the variable, x. At least, the errors in x are assumed to be negligible in comparison with those of y.[1]

A final word of caution must be added with regard to the *use* of the empirical formula. In general, it is safe to use it within the range of the observations, that is, in interpolation; but only in very exceptional cases should it be depended upon for extrapolation, outside of these limits. Duncan [2] cites the example of the stress-strain curve, which is practically a straight line until the elastic limit is reached, but which, at that point, suddenly breaks into a sharp curve. An extrapolation from the straight line would be greatly in error.

Again, it must be emphasized that the *form* of the empirical equation is *assumed* at the outset and from considerations outside

[1] For an investigation of the case when x and y are equally subject to error, see Report of C. & G. Survey, 1890, page 687, or Wright's Adjustment of Observations, Art. 106.

[2] Practical Curve Tracing by R. H. Duncan.

of the Method of Least Squares. From that point as a beginning, this method determines the best values of the coefficients *for that form of equation*, but unless a suitable form has been selected the resulting empirical formula may be no better than a rough guess. Therefore, great care should be exercised in choosing the form of the equation.

When the observed data are few and widely scattered, it is scarcely worth while to go to the trouble of a Least Squares adjustment to establish an empirical formula. In such a case, it is usually sufficient to sketch a smooth curve through the plotted observations and to determine the constants of the curve by scaling various elements from it, in connection with its known properties. In particular is this method applicable to straight lines and to those hyperbolic forms which appear as straight lines when plotted on logarithmic paper.

125. Example: Straight Line. Let it be required to derive a formula which shall fit the following observations as nearly as possible, preference being given to a straight line, if reasonable.

x	y	x	y
−1.0	+14.0	+14.0	+5.0
+1.0	13.0	17.0	2.9
5.0	10.7	20.0	1.0
9.0	8.0		

Upon plotting these observations, as in Fig. 34, it is seen that they fall nearly in a straight line, so we shall assume the form

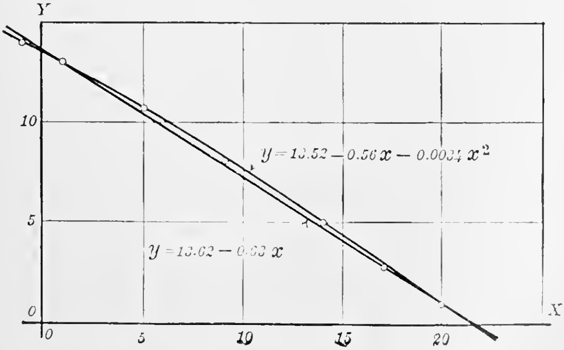


FIG. 34. Straight Line and Parabola

of the equation to be, $y=A+Bx$, A and B to be determined. Substituting the observed data in this form and reversing the order we obtain the equations,

$$
\begin{aligned}
-\ \ \ B+A-14.0&=0\\
+\ \ \ B+A-13.0&=0\\
+\ 5B+A-10.7&=0\\
+\ 9B+A-\ \ 8.0&=0\\
+14B+A-\ \ 5.0&=0\\
+17B+A-\ \ 2.9&=0\\
+20B+A-\ \ 1.0&=0
\end{aligned}
\tag{155}
$$

Considering these to have the form, $a_1A+b_1B+l_1=v_1$, to correspond to equations (18), the normal equations take the form of (22), and become

$$
\begin{aligned}
+993B+65A-263.8&=0\\
+\ 65B+\ 7A-\ 54.6&=0
\end{aligned}
\tag{156}
$$

the solution of which gives $A=+13.62$, and $B=-0.63$, so that the required empirical formula is,

$$y=13.62-0.63x \tag{157}$$

Substituting the values of A and B in (155), with the original values of x, the computed values, y', are obtained, and subtracting from these the corresponding observed values, y, we find the residuals v, which, for further reference, will be squared and added.

x	y	y'	v	v^2
-1.0	$+14.0$	$+14.25$	$+.25$	.0625
$+1.0$	13.0	12.99	$-.01$	1
5.0	10.7	10.47	$-.23$	529
9.0	8.0	7.95	$-.05$	25
14.0	5.0	4.80	$-.20$	400
17.0	2.9	2.91	$+.01$	1
20.0	1.0	1.02	$+.02$	4
				$[v^2]=.1585$
0.0		$+13.62$		
$+21.6$		0.0		

The line is easily plotted from the points where it crosses the axes, that is, where $x=0$ and where $y=0$, which have been added to the above table. It is shown in Fig. 34. The residuals are indicated as the vertical distances of the observations from the plotted line.

126. Example: Parabola. From the observations in the preceding article, let us determine a curve instead of a straight line, using the parabolic form,

$$y=A+Bx+Cx^2 \tag{158}$$

Substituting the observed values of x and y, and reversing the order, we obtain the observation equations,

$$\begin{aligned}
C-\quad B+A-14.0&=0\\
C+\quad B+A-13.0&=0\\
25C+\ 5B+A-10.7&=0\\
81C+\ 9B+A-\ \ 8.0&=0\\
196C+14B+A-\ \ 5.0&=0\\
289C+17B+A-\ \ 2.9&=0\\
400C+20B+A-\ \ 1.0&=0
\end{aligned} \tag{159}$$

In order to reduce the coefficients of the first two unknowns, we let $C'=100C$, and $B'=10B$, as in Art. 49. Then we have,

$$\begin{aligned}
.01C'-\ .1B'+A-14.0&=0\\
.01C'+\ .1B'+A-13.0&\\
.25C'+\ .5B'+A-10.7&\\
.81C'+\ .9B'+A-\ \ 8.0&\\
1.96C'+1.4B'+A-\ \ 5.0&\\
2.89C'+1.7B'+A-\ \ 2.9&\\
4.00C'+2.0B'+A-\ \ 1.0&
\end{aligned} \tag{160}$$

The resulting normal equations are,

C'	B'	A		
+28.91	+16.50	+9.93	−31.61	=0
+16.50	+ 9.93	+6.50	−26.38	
+ 9.93	+ 6.50	+7.00	−54.60	

(161)

and their solution yields the values, $A=+13.52$, $B'=-5.63$, and $C'=-0.34$. Then $B=0.1B'=-0.56$, and $C=0.01C'=-0.0034$. The empirical formula is, therefore,

$$y=13.52-0.56x-0.0034x^2 \tag{162}$$

The similarity of the first two terms of the second member to those of (157), as well as the very small coefficient of x^2, indicates that the curve approximates closely to the straight line of the previous article. However, we shall investigate the residuals to see how closely the observations are followed. Computing the values of y and designating them by y' as before, we find:

x	y	y'	v	v^2
− 1.0	+14.0	+14.08	+0.08	.0064
+ 1.0	13.0	12.96	− 4	16
5.0	10.7	10.64	− 6	36
9.0	8.0	8.20	+ 20	400
14.0	5.0	5.01	+ 1	1
17.0	2.9	3.02	+ 12	144
20.0	1.0	0.96	− 4	16
			$[v^2]=$	.0677

Evidently, this curve is much closer to the observations than is the straight line. The residuals are smaller and the sum of their squares is smaller by more than half. The plotted curve is shown, also, in Fig. 34, where its advantages are apparent.

127. Example: Exponential Curve. The following observations are plotted in Fig. 35, and an exponential curve seems reasonable to assume. In order to investigate the equation more in

x	y	$\log x$	diff.	$\log y$	diff.	y^2
0.2	4.6	9.301		0.663		21
0.6	5.8	9.778	.48	0.763	.10	34
1.2	7.6	0.079	.30	0.881	.12	58
1.6	9.6	0.204	.12	0.982	.10	92
2.0	11.5	0.301	.10	1.061	.08	132
2.4	14.4	0.380	.08	1.158	.10	207
2.8	17.5	0.447	.07	1.243	.08	306
3.0	20.0	0.477	.03	1.301	.06	400

detail, the common logarithms of x and y are tabulated, also, with their successive differences.

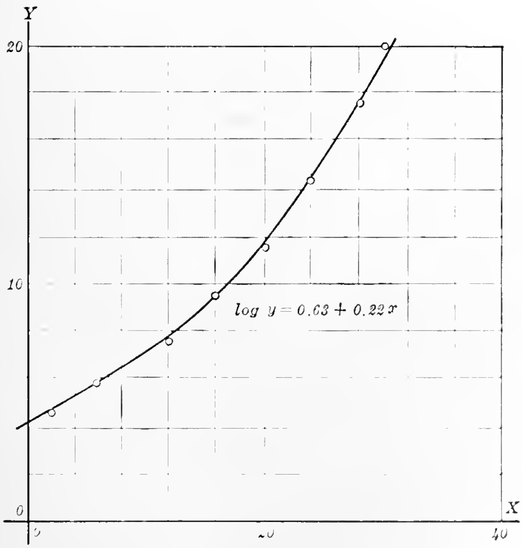


FIG. 35. Exponential Function; Simple Plotting

It is evident from an inspection of these differences that there is no straight-line relation between $\log x$ and $\log y$, and plotting these values as coordinates shows a distinct curve, in Fig. 36. However, the differences in log y are seen to correspond quite

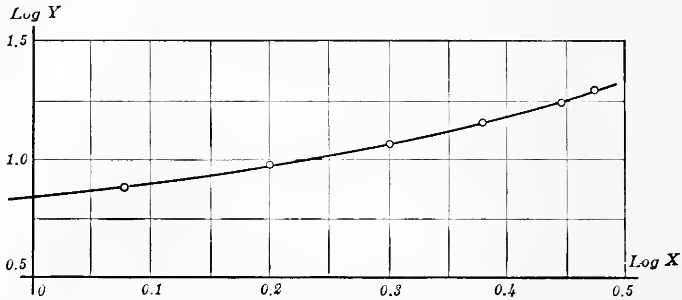


FIG. 36. Exponential Function; Logarithmic Plotting

closely with those in x itself, and this is verified by plotting x and log y, in Fig. 37. Therefore, the desired equation will have

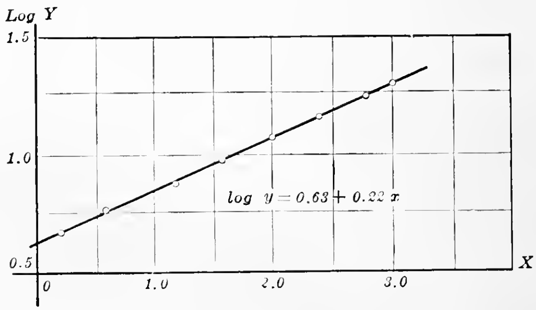


FIG. 37. Exponential Function; Semi-logarithmic Plotting

the form,

$$\log y = A + Bx \tag{163}$$

or, if $A = \log A'$,

$$y = A'(10^{Bx}) \tag{164}$$

Writing the former equation in the usual order for observation equations,

$$Bx + A - \log y = 0 \tag{165}$$

Substituting the values of x and log y from the table, above, and carrying the numerical work to two places, only, we have,

$$\begin{array}{ll} & \text{Wt.} \\ 0.2B+A-0.66=0 & 0.2 \\ 0.6B+A-0.76 & 0.3 \\ 1.2B+A-0.88 & 0.6 \\ 1.6B+A-0.98 & 0.9 \\ 2.0B+A-1.06 & 1.3 \\ 2.4B+A-1.16 & 2.1 \\ 2.8B+A-1.24 & 3.1 \\ 3.0B+A-1.30 & 4.0 \end{array} \tag{166}$$

The weights of the original observations are assumed equal. Those of log y, and the observation equations, will then be directly as the squares of the y's. In the table these have been divided by 100 to lessen numerical labor.

The normal equations, formed in accordance with (21), are,

$$\begin{aligned} +80.88B+30.70A-37.18&=0 \\ +30.70B+12.50A-14.63&=0 \end{aligned} \tag{167}$$

and from their solution, $B=+0.22$ and $A=+0.63$, so that the empirical formula is,

$$\log y=0.63+0.22x \tag{168}$$

or,

$$x=4.55 \log y-2.86 \tag{169}$$

or,

$$y=4.27(10^{0.22x}) \tag{170}$$

Computing the values of log y' and from them those of y', corresponding to the successive values of x, and forming the residuals, we obtain the following table:

x	log y'	log y	r_1	y'	y	r_2
0.2	.67	.66	+.01	4.7	4.6	+0.1
0.6	.76	.76	0	5.8	5.8	0
1.2	.89	.88	1	7.8	7.6	2
1.6	.98	.98	0	9.6	9.6	0
2.0	1.07	1.06	1	11.8	11.5	3
2.4	1.16	1.16	0	14.4	14.4	0
2.8	1.25	1.24	+.01	17.8	17.5	+0.3
3.0	1.29	1.30	−.01	19.5	20.0	−0.5

The curve is plotted in Fig. 35, and the straight line, using log y, in Fig. 37. The residuals of log y, in the column headed v_1, are practically negligible; those of y, called v_2, are somewhat larger, and increase numerically with x. This may seem surprising in view of the increasing weight used, and in order to illustrate this effect, the normal equations were formed a second time without considering weights at all, and solved with the following equation as a result:

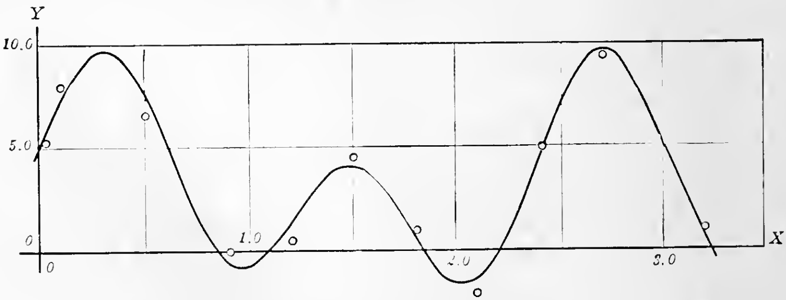
$$\log y = +0.61 + 0.23x \tag{171}$$

The residuals of log y are about the same as before, but those of y are, respectively, -1, 0, 0, 0, $+.3$, 0, $+.7$, and 0, indicating the diminishing weights. However, the curve follows the observations so closely that the weights have little effect upon the empirical formula.

128. Example: Periodic Curve. The following set of observations is given for the purpose of determining the equation which will best represent them. They are of equal weight.

x	y	x	y
+0.1	+8.0	+1.8	+1.0
0.5	+6.8	2.1	−2.0
0.9	0.0	2.4	+5.0
1.2	+0.5	2.7	+9.5
1.5	+4.5	3.2	+1.0

These data are plotted in Fig. 38 and from a curve sketched through the points it is evident that the function is a periodic one. Also



$y = 3.00 - 2.24 \sin 150x^\circ + 1.73 \cos 150x^\circ + 3.93 \sin 300x^\circ + 0.09 \cos 300x^\circ$

Fig. 38. Compound Periodic Curve

the waves occur in pairs, one large and the next smaller. The cycle or period is completed in approximately 2.4 units of x, and this value will be assumed for m in (143). Owing to the fact that the waves in the curve are not equal, the first five terms of (143) will be used, namely,

$$y=A+B\sin\frac{360^\circ}{m}x+C\cos\frac{360^\circ}{m}x$$

$$+D\sin\frac{360^\circ}{m}2x+E\cos\frac{360^\circ}{m}2x \qquad (172)$$

which becomes, upon inserting the above value of m,

$$y=A+B\sin 150x^\circ+C\cos 150x^\circ$$

$$+D\sin 300x^\circ+E\cos 300x^\circ \qquad (173)$$

Substituting the various values of x and y, and looking up the natural sines and cosines to two places, we obtain the observation equations, which will be written, for convenience, in the reverse order:

$$\begin{aligned}
&+0.87E+0.50D+0.97C+0.26B+A-8.0=0\\
&-0.87E+0.50D+0.26C+0.97B+A-6.8=0\\
&\qquad\quad -1.00D-0.71C+0.71B+A\qquad=0\\
&+1.00E\qquad\qquad -1.00C\qquad\qquad +A-0.5=0\\
&\qquad\quad +1.00D-0.71C-0.71B+A-4.5=0\\
&-1.00E\qquad\qquad\qquad\qquad -1.00B+A-1.0=0\\
&\qquad\quad -1.00D+0.71C-0.71B+A+2.0=0\\
&+1.00E\qquad\qquad +1.00C\qquad\qquad +A-5.0=0\\
&\qquad\quad +1.00D+0.71C+0.71B+A-9.5=0\\
&-0.50E-0.87D-0.50C+0.87B+A-1.0=0
\end{aligned} \qquad (174)$$

The normal equations, formed in the usual manner, are,

E	D	C	B	A			
+4.77	+0.44	+0.86	−0.05	+ 0.50	− 5.04 = 0		
	+5.26	+1.05	−0.15	+ 0.13	−22.53 = 0		
		+5.26	+0.06	+ 0.73	−15.64 = 0	(175)	
			+4.77	+ 1.10	−13.51 = 0		
				+10.00	−34.30 = 0		

the sub-diagonal terms being omitted for the abridged solution. Solving these equations, the following values of the unknowns are obtained:

$A=+3.00,\quad B=+2.24,\quad C=+1.73,\quad D=+3.93,\quad E=+0.09.$

The empirical formula, therefore, will be,

$$\begin{aligned} y=3.00&+2.24 \sin 150x^\circ+1.73 \cos 150x^\circ \\ &+3.93 \sin 300x^\circ+0.09 \cos 300x^\circ \end{aligned} \tag{176}$$

or, expressed in the form mentioned at the close of Art. 118,

$$\begin{aligned} y=3.00&+2.83 \cos (150x^\circ-52^\circ\ 19') \\ &+3.93 \cos (300x^\circ-88^\circ\ 41') \end{aligned} \tag{177}$$

The curve is plotted in Fig. 38. For this purpose, a number of extra values of y were computed so as to determine the curve with greater precision. It is evident that a larger number of observations would be desirable in the case of an equation as complicated as this one. The curve conforms to the observations fairly well, and it is doubtful that a recomputation with a different value of m for the period would be worth while. It is useful to note in connection with the plotting that the same value of y will correspond to values of x which differ by multiples of m. Thus, for $x=0.1$ and 2.5, we have the same value of y, namely, $+7.32$.

129. References. The reader is referred to the following works in which useful information and methods concerning empirical formulas will be found. The collections of examples given by Weld and Bartlett are worthy of note.

Wright: Adjustment of Observations.
Comstock: Method of Least Squares.
Merriman: Method of Least Squares.
Weld: Theory of Errors and Least Squares.
Bartlett: Method of Least Squares.
Helmert: Ausgleichungsrechnung.
Duncan: Practical Curve Tracing.
Brunt: Combination of Observations.
Lipka: Graphical and Mechanical Computation.

CHAPTER VIII

PRECISION OF OBSERVATIONS AND RESULTS AND COMBINATION OF COMPUTED QUANTITIES

130. Having considered the determination of the best values of the unknown quantities to be obtained from given observations, it remains to investigate the degree of confidence which may be placed in the observations and the computed results, so that they may be compared with the results of other observations of the same quantities.

131. Precision. If two sets of direct observations of the same kind be compared, and in the first the component quantities are scattered over a wider range or are more discordant than in the second, it is natural to conclude that the observations of the first set were made with less care or under less favorable circumstances than those of the second set. The latter are more consistent and evidently more precise; their differences or discrepancies are smaller. Furthermore, even though the number of observations in the two sets were equal, the mean of the second set would be regarded as of greater reliability or weight than the mean of the first set, merely because of the greater consistency, i.e., smaller discrepancies, among its original observations. Since these smaller discrepancies correspond to smaller residuals from the mean, it is evident that the precision of the mean is indicated by the size of its residuals, being greater as the residuals are smaller, and vice versa.

132. Precision and Accuracy.[1] This precision must not be confused with the accuracy, that is, the correctness, of the results. The latter is affected by systematic errors (Art. 5). Thus, a series of observations may be very closely grouped, showing a

[1] See Johnson, Theory of Errors and Method of Least Squares, Chap. VII, for an extended treatment of this subject.

high degree of precision, but each separate observation, and therefore, the mean, may be in error by a large amount due to some influence which is unknown or not taken into account. Precision has reference to the accidental errors of observations made under constant conditions and indicates the care exercised by the observer, the closeness with which the instrumental readings are made, and the suitability of the method used. Discordant observations are not precise; but precise determinations may or may not be accurate.

133. Index of the Precision. It is easy to obtain an idea as to the precision of the observations from an inspection of them or of the residuals of their mean. But in the comparison of the results of different sets of observations of the same quantities, it is very convenient to have a numerical index from which the precision of each set may be determined without actually inspecting the observations themselves. Since this precision is indicated, in general, by the size of the residuals, it is evident that the desired index would logically be some function of these residuals. The precision of a result will depend, also, upon the number of observations from which it is obtained. Obviously, the larger the series of observations, the greater should be the precision of their mean as well as that of the typical single observation. Thus, we might use the mean of the residuals, without regard to signs, or the square root of the sum of their squares, and either of these would give us some idea of the consistency of the observations, this hypothetical residual being smaller in the case of greater precision.

From the very inception of the Method of Least Squares, the investigation of the precision was regarded as of considerable importance. Several quantities have been used to indicate it. Gauss designated the quantity, h, in the Law of Error, as a "measure of precision." However, other indices have been more generally used, namely, certain selected errors, theoretically defined, as the *Mean Square Error*, the *Probable Error*, and the *Average Error*. These will now be considered in order.

134. The Quantity, h, in the Law of Error. If we consider two sets of observations of the same quantity, made in the same manner, to be represented by the curves in Fig. 39, the area between each curve and the axis of Δ will be unity, that is, the probability of an error between the limits $-\infty$ and $+\infty$. Then, the taller the curve, i.e., the greater the p-intercept, the larger will be the portion of the area immediately adjacent to the p-axis and the more numerous the smaller errors will be in comparison with the entire group; in other words, the greater will be the precision. By inspection of the Law of Error,

$$p=\frac{h}{\sqrt{\pi}}e^{-h^2\Delta^2}$$

it is seen that when $\Delta=0$, $p=\frac{h}{\sqrt{\pi}}$, so that the p-intercept is $\frac{h}{\sqrt{\pi}}$. Therefore, $\sqrt{\pi}$ being a constant, h may be regarded as an index of the precision, with which it varies directly.

FIG. 39. Curves of Probability

135. The Mean Square Error (ϵ) of an observation is defined as the square root of the mean of the squares of the errors in a given series of observations.[1] It will be represented by ϵ or m. s. e. To determine its relation to h of the previous article, we proceed as follows:

According to the Law of Error, the probability of the occurrence of an error, Δ, in a given set of observations, is,

$$p=f(\Delta)=\frac{h}{\sqrt{\pi}}e^{-h^2\Delta^2} \qquad (178)$$

[1] The mean square error is sometimes referred to as the *mean error*. This introduces an ambiguity with the *average error*, or mean of the errors, and is an unfortunate use of the term. German writers call it *der mittlere Fehler* but this involves no ambiguity as they designate the average error as *der durchschnittliche Fehler*.

and the probability of an error between the limits Δ and $\Delta+d\Delta$ is,

$$\frac{h}{\sqrt{\pi}}e^{-h^2\Delta^2}d\Delta \tag{179}$$

The number of these errors will be equal to their probability times the total number of observations (or errors) in the set,[1] that is,

$$\frac{nh}{\sqrt{\pi}}e^{-h^2\Delta^2}d\Delta \tag{180}$$

and the sum of their squares will be,

$$\frac{nh}{\sqrt{\pi}}e^{-h^2\Delta^2}\Delta^2 d\Delta \tag{181}$$

then the sum of the squares of all of the errors, between the limits $-\infty$ and $+\infty$, will be,

$$\frac{nh}{\sqrt{\pi}}\int_{-\infty}^{+\infty}e^{-h^2\Delta^2}\Delta^2 d\Delta \tag{182}$$

and the mean of their squares, equal to ϵ^2 by definition, is

$$\epsilon^2=\frac{nh}{n\sqrt{\pi}}\int_{-\infty}^{+\infty}e^{-h^2\Delta^2}\Delta^2 d\Delta \tag{183}$$

$$=\frac{h}{\sqrt{\pi}}\int_{-\infty}^{+\infty}e^{-h^2\Delta^2}\Delta^2 d\Delta \tag{184}$$

Substituting in (184) the value of the definite integral,[2]

[1] See Appendix C.

[2] This integral may be evaluated in the following manner (Bartlett): The probability of an error between the limits $-\infty$ and $+\infty$ is unity (certainty), that is,

$$\frac{h}{\sqrt{\pi}}\int_{-\infty}^{+\infty}e^{-h^2\Delta^2}d\Delta=1 \tag{184a}$$

or,

$$\int_{-\infty}^{+\infty}e^{-h^2\Delta^2}d\Delta=\frac{\sqrt{\pi}}{h} \tag{184b}$$

Differentiating both members with respect to h, we obtain,

$$-2h\int_{-\infty}^{+\infty}e^{-h^2\Delta^2}\Delta^2 d\Delta dh=-\frac{\sqrt{\pi}}{h^2}dh \tag{184c}$$

whence,

$$\int_{-\infty}^{+\infty}e^{-h^2\Delta^2}\Delta^2 d\Delta=\frac{\sqrt{\pi}}{2h^3} \tag{184d}$$

For another solution, see Jordan, Handbuch der Vermessungskunde, I, 561.

$$\epsilon^2 = \frac{h}{\sqrt{\pi}}\left(\frac{\sqrt{\pi}}{2h^3}\right) = \frac{1}{2h^2} \tag{185}$$

Hence,

$$\epsilon = \frac{1}{h\sqrt{2}} \tag{186}$$

or,

$$h = \frac{1}{\epsilon\sqrt{2}}, \text{ as stated in Art. 19.}$$

The geometrical interpretation of the mean square error is that it corresponds to the abscissa of the point of inflexion of the Error Curve. Differentiating (178) and placing the second derivative equal to zero, we have,

$$f'(\Delta) = \frac{-2h^3\Delta}{\sqrt{\pi}}e^{-h^2\Delta^2} \tag{187}$$

$$f''(\Delta) = \frac{-2h^3}{\sqrt{\pi}}e^{-h^2\Delta^2} + \frac{4h^5\Delta^2}{\sqrt{\pi}}e^{-h^2\Delta^2} \tag{188}$$

$$= \frac{2h^3}{\sqrt{\pi}}e^{-h^2\Delta^2}(2h^2\Delta^2 - 1) = 0 \tag{189}$$

Therefore, for the point of inflexion,

$$2h^2\Delta^2 - 1 = 0 \quad \text{or,} \quad 2h^2\Delta^2 = 1 \tag{190}$$

and

$$\Delta = \frac{1}{h\sqrt{2}} = \epsilon, \quad \text{from (186)} \tag{191}$$

which shows that the point of inflexion corresponds to the mean square error of an observation.

136. The Probable Error (r) of an observation in a given series is the middle one of all the errors when they are arranged in numerical order, each being written as many times as it occurs. As many of the errors are greater than it as are less, and so the probability of an error greater than the probable error is equal to that of an error less than it, namely, 0.5, since the total probability is unity. It is an even chance that an error taken at random from the series will be greater or less than the probable error.

This is not the *most probable error* in the series, for that would be zero, to correspond to the maximum ordinate to the Error Curve, and it is unfortunate that the name has come to be quite generally used in this country. It is simply a quantity from which the precision of the observations can be estimated or determined, in comparison with similar quantities referring to other observations. A better name for it would be the *middle error.* It is represented by the letter r.

The probability that the error of an observation will be numerically less than the probable error is, by definition, $\frac{1}{2}$. Then from the law of Error,

$$\frac{h}{\sqrt{\pi}}\int_{-r}^{+r} e^{-h^2\Delta^2}d\Delta=\frac{1}{2} \tag{192}$$

or changing the lower limit,

$$\frac{2h}{\sqrt{\pi}}\int_{0}^{r} e^{-h^2\Delta^2}d\Delta=\frac{1}{2} \tag{193}$$

It is not feasible to determine the value of r in terms of h directly from this equation, so we make use of the following process:

In the Law of Error, let

$$t=h\Delta, \quad \text{whence } d\Delta=\frac{dt}{h}.$$

Then we have for the probability of an error less than Δ,

$$\frac{2h}{\sqrt{\pi}}\int_{0}^{\Delta} e^{-h^2\Delta^2}d\Delta=\frac{2}{\sqrt{\pi}}\int_{0}^{t} e^{-t^2}dt \tag{194}$$

This expression is evaluated for various values of t, by expansion into a series,[1] and the results are tabulated with t as an argument.[2] By interpolation in this table with the value of the probability 0.5, the corresponding value of t is found to be 0.4769, which is the value of $t=h\Delta$ when Δ is the probable error, r. Thus,

$$hr=0.4769 \tag{195}$$

and

$$r=\frac{0.4769}{h} \tag{196}$$

Since the probability that an error will lie between certain limits is represented by the *area* bounded by the Error Curve, the horizontal axis, and the ordinates at those limits; and since

[1] See Appendix C, page 215.

[2] See Table I, page 229.

the entire area between the curve and the horizontal axis represents the probability of an error between $-\infty$ and $+\infty$, that is, unity (certainty); it follows that the ordinate of the probable error divides the area on either side of the vertical axis into two equal parts corresponding to the probability, $\frac{1}{2}$.

137. The Average Error (η) is the mean of all the errors in a set without regard to signs. Since positive and negative errors are equally likely to occur, the probability of a *positive* error between Δ and $\Delta+d\Delta$ will be one-half of that of *any* error between those limits, that is, it will be equal to

$$\frac{h}{2\sqrt{\pi}}\int_{-\infty}^{+\infty} e^{-h^2\Delta^2}d\Delta \tag{197}$$

The number of the positive errors will be their probability times the total number of errors, n, namely,

$$\frac{nh}{2\sqrt{\pi}}\int_{-\infty}^{+\infty} e^{-h^2\Delta^2}d\Delta \tag{198}$$

and their sum is,

$$\frac{nh}{2\sqrt{\pi}}\int_{-\infty}^{+\infty} e^{-h^2\Delta^2}\,\Delta d\Delta. \tag{199}$$

But the sum of the negative errors is numerically equal to that of the positive ones, so that the total sum will be twice the above, that is,

$$\frac{nh}{\sqrt{\pi}}\int_{-\infty}^{+\infty} e^{-h^2\Delta^2}\,\Delta d\Delta. \tag{200}$$

or,

$$\frac{2nh}{\sqrt{\pi}}\int_{0}^{\infty} e^{-h^2\Delta^2}\,\Delta d\Delta. \tag{201}$$

and the average of all of the errors is therefore,

$$\eta=\frac{2h}{\sqrt{\pi}}\int_{0}^{\infty} e^{-h^2\Delta^2}\Delta d\Delta \tag{202}$$

$$=\frac{-1}{h\sqrt{\pi}}\int_{0}^{\infty} e^{-h^2\Delta^2}(-2h^2\Delta)d\Delta \tag{203}$$

$$=\frac{-1}{h\sqrt{\pi}}\Big[e^{-h^2\Delta^2}\Big]_0^{\infty} \tag{204}$$

so that,

$$\eta=\frac{1}{h\sqrt{\pi}} \tag{205}$$

The ordinate of the average error passes through the center of gravity of the area between the curve of error and the horizontal axis on either side of the vertical axis. For, if Δ_0 represent the abscissa of the center of gravity, by considering vertical strips of width $d\Delta$ and length

$$\frac{h}{\sqrt{\pi}}e^{-h^2\Delta^2}$$

and taking moments about the origin, we have,

$$\Delta_0\frac{h}{\sqrt{\pi}}\int_0^\infty e^{-h^2\Delta^2}d\Delta=\frac{h}{\sqrt{\pi}}\int_0^\infty e^{-h^2\Delta^2}\Delta d\Delta \qquad (206)$$

But since the total probability area is equal to unity, the area on one side of the vertical axis is 1/2, that is

$$\frac{h}{\sqrt{\pi}}\int_0^\infty e^{-h^2\Delta^2}d\Delta=\tfrac{1}{2} \qquad (207)$$

Hence,

$$\Delta_0=\frac{2h}{\sqrt{\pi}}\int_0^\infty e^{-h^2\Delta^2}\Delta d\Delta \quad =\eta, \text{ from (202).} \qquad (208)$$

138. Comparison of the Indices of Precision. From (186), (196), and (205), we obtain directly,

$$h=\frac{1}{\epsilon\sqrt{2}}=\frac{0.4769}{r}=\frac{1}{\eta\sqrt{\pi}} \qquad (209)$$

$$\frac{1}{h}=1.4142\epsilon=2.0966r=1.7726\eta \qquad (210)$$

$$\left.\begin{aligned}\epsilon&=1.4826r=1.2533\eta\\ r&=0.6745\epsilon=0.8453\eta\\ \eta&=0.7979\epsilon=1.1829r\end{aligned}\right\} \qquad (211)$$

Thus it is seen that the mean square error, the probable error, and the average error are related by constant factors. Therefore, they may be used interchangeably in various formulas and mathematical investigations by simply providing for the numerical factors.

In Fig. 40, these quantities are shown in their correct relative positions and magnitudes. The abscissæ represent the errors and the ordinates their probabilities. It will be recalled that the intercept on the vertical axis is $\frac{h}{\sqrt{\pi}}$.

The quantity, h, is directly proportional to the precision. However, it is inconvenient in practice and is not generally used. The three representative errors, ϵ, r, and η, on the other hand, are inversely proportional to the precision; the smaller these errors, the more precise and consistent are the observations. They are sometimes said to indicate the *uncertainty*, therefore, instead of the precision. Each of the three errors occurs in a

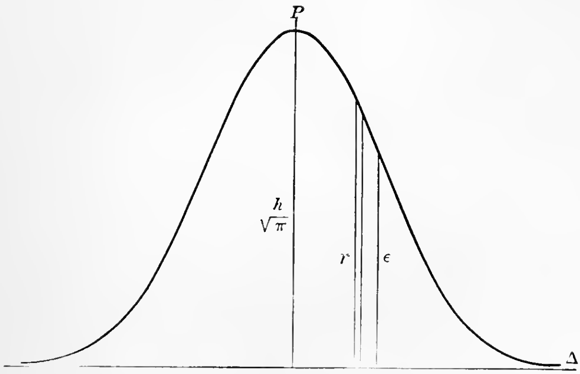


FIG. 40. Relations between the Various Indices of Precision

certain relative position when all the errors in a set of observations are arranged in the order of their numerical magnitude, as, for example, the probable error occupies the middle of the series. This feature is what one would naturally expect in an index of the precision (Art. 133).

The average error, also, is not used in practice as an index, although it would be a satisfactory one. It may be used, however, in the process of determining ϵ and r.

The mean square error and the probable error are in common use as indices of the precision. The former has been almost universally used by writers in German and other foreign languages, as well as by some Americans, notably in the classic Adjustment of Observations, by Professor T. W. Wright, and by Chauvenet, Newcomb, and Crandall. Its principal advantages lie in the facility of its theoretical derivation; in its priority (it was used by Gauss); in its use by the Germans and French, who have made the most numerous contributions to the subject of Least Squares;[1] and in its avoidance of the misnomer of the probable error which is frequently a stumbling-block to the beginner.

The probable error is used by most American and British writers. Its name is its greatest enemy, but there may be some advantage in its mere reference to probability. The person who does not clearly understand its significance is apt to take it at its face value and so interpret it. However, it is hoped that such persons will learn what it means or leave it alone. It should be understood simply as an index of the precision.

Whichever index is used, it is written after the quantity to which it refers and separated from it by the sign, $\pm$. This is merely a convention and the sign is never to be used algebraically. There is never any reason for increasing or diminishing a quantity by the amount of its mean square error or probable error. A better method of designating it would be to use instead of the sign, $\pm$, the symbol for the mean square error (ϵ or m. s. e.), or that for the probable error (r or p. e.), as 7653.28 ($\epsilon=0.02$). But the use of the $\pm$ sign is well established as is also the term probable error.

139. Precision of Direct Observations. We have seen how the precision in a set of observations may be indicated by the mean square error, the probable error, etc., and it is evident that if we could know the true value of the observed quantity, and therefore, the true errors, Δ, we could ascertain the numerical value of the index of the precision directly from these errors, by definition, as

$$\epsilon^2=\frac{[\Delta^2]}{n}, \quad \eta=\frac{[\Delta}{n}, \text{ and } r=\text{the middle error.}^2$$

But as these true errors are unknown, it remains to determine the precision index from the given observations or residuals. Know-

[1] See Appendix A.

[2] The symbol for the sum without regard to signs is [

ing the relations between the three indices as stated in (211), it will suffice to determine the mean square error in each case and from it to express the probable error and the average error.

140. Precision of a Single Observation. Each observation has its own individual error and when we refer to a "single observation" in this connection, we mean an observation such as those in the set which is being discussed, not any single one of them, but a hypothetical one which is never evaluated, but which is typical of the entire set in so far as precision is concerned.

Using the notation of Chapter I, let M represent an observation of a directly observed quantity; v, its residual from the arithmetic mean, x_0; X, the true value of the observed quantity; Δ, the true error of an observation; Δ_0, the true error of the arithmetic mean; and n, the number of observations in the series. Then, using subscripts to indicate the separate observations,

$$X = x_0 + \Delta_0 = M_1 + \Delta_1 = M_2 + \Delta_2 \quad . \quad . \quad . \tag{212}$$

$$v_1 = x_0 - M_1, \quad v_2 = x_0 - M_2 \quad . \quad . \quad . \tag{213}$$

$$\Delta_1 = x_0 + \Delta_0 - M_1, \quad \Delta_2 = x_0 + \Delta_0 - M_2, \quad . \quad . \quad . \tag{214}$$

$$\Delta_1 = v_1 + \Delta_0, \qquad \Delta_2 = v_2 + \Delta_0, \quad . \quad . \quad . \tag{215}$$

Squaring both members and adding the n resulting equations,

$$[\Delta^2] = [v^2] + 2\Delta_0[v] + n\Delta_0^2 \tag{216}$$

From (8), $[v] = 0$; and the unknown true error, Δ_0, of the mean is assumed, for this demonstration, to be equal to the mean square error of the mean, the value of which is determined in Art. 153 to be (see Art. 141),

$$\epsilon_0 = \frac{\epsilon}{\sqrt{n}}.$$

Therefore, dividing (216) by n,

$$\frac{[\Delta^2]}{n} = \epsilon^2 = \frac{[v^2]}{n} + \frac{\epsilon^2}{n} \tag{217}$$

whence,

$$\epsilon^2 = \frac{[v^2]}{n-1} \tag{218}$$

and

$$\epsilon = \sqrt{\frac{[v^2]}{n-1}} \tag{219}$$

Then, from (211),

$$r = 0.6745\sqrt{\frac{[v^2]}{n-1}} \tag{220}$$

$$\eta = 0.7979\sqrt{\frac{[v^2]}{n-1}} \tag{221}$$

These three formulas are known as Bessel's Formulas and the first two are in general use. In long series of observations, however, it is more convenient to use Peters' Formulas, which involve the sum of the residuals without regard to signs, $[v$, instead of the sum of their squares. They may be derived as follows:

From (217) and (218),

$$\frac{[\Delta^2]}{n} = \frac{[v^2]}{n-1} \tag{222}$$

$$[v^2] = \frac{n-1}{n}[\Delta^2] \tag{223}$$

and,

$$v_1 = \sqrt{\frac{n-1}{n}}\Delta_1, \quad v_2 = \sqrt{\frac{n-1}{n}}\Delta_2, \ . \ . \ . \tag{224}$$

Adding these n equations, neglecting the signs of v and Δ, we have, since, by definition, $\eta = [\Delta/n$,

$$[v = \sqrt{\frac{n-1}{n}}[\Delta = n\sqrt{\frac{n-1}{n}}\eta \tag{225}$$

whence,

$$\eta = \frac{[v}{\sqrt{n(n-1)}} \tag{226}$$

and from (211),

$$\epsilon = 1.2533\frac{[v}{\sqrt{n(n-1)}} \tag{227}$$

$$r = 0.8453\frac{[v}{\sqrt{n(n-1)}} \tag{228}$$

An approximate value of the probable error of a single observation in a series of from 20 to 30 observations may be determined by taking one-sixth of the range of the set, or one-third of the largest residual. From the table of values of the Law of Error, it is found that the probability of an error three times as great as the probable error is about 0.04, or 1 in 25.[1] That is, in a series of 25 observations, the maximum error is likely to be about three times the probable error of a single observation. And, since there are as many positive as negative errors, the total range of the observations in an ordinary set of, say, from 20 to 30 observations, is likely to be about six times the probable error, or about four times the mean square error of a single observation. Conversely, knowing the precision index, and the approximate number of observations in the set, we can estimate the range. Frequently this fact affords the most tangible idea as to the consistency of the observations, especially to the beginner, since, by doubling the mean square error of a single observation he obtains an approximate value of the maximum residual.

141. Precision of the Mean. The arithmetic mean being the best value of the observed quantity obtainable from the given direct observations (Arts. 14, 27), it is obvious that the precision of the mean will be greater than that of a single observation, and also that the precision will increase with the number of observations in the set. In Art. 153 it is shown that if ϵ be the mean square error of a single observation, and ϵ_0 that of the mean of the set of n observations,

$$\epsilon_0 = \frac{\epsilon}{\sqrt{n}} \tag{229}$$

which expresses the very important relation that the *precision of the mean increases directly as the square root of the number of observations.* In other words, to double the precision, that is, to divide ϵ_0 by two, it is necessary to make four times as many observations.[2]

[1] See Appendix F, page 231.

The probability of an error less than three times the probable error is 0.957, corresponding to $\Delta\ r = 3.0$; then the probability of an error greater than this would be $1 - 0.957 = 0.043$.

[2] It must not be assumed that by increasing the number of observations without limit, the precision can be indefinitely increased. There are always influences which make it extremely difficult, if not quite impossible, to approach certainty beyond a definite limit. In this connection, the reader is referred to the admirable treatment of this matter in Wright and Hayford's Adjustment of Observations, Arts. 38 to 46.

This principle is used in determining the most economical or advisable number of observations to make in a certain program.

From (229) and the formulas of the preceding article, we obtain the following expressions for the three precision indices of the mean, by dividing by $\sqrt{n}$ in each case;

Bessel's Formulas:

$$\epsilon_0=\sqrt{\frac{[v^2]}{n(n-1)}} \tag{230}$$

$$r_0=0.6745\sqrt{\frac{[v^2]}{n(n-1)}} \tag{231}$$

$$\eta_0=0.7979\sqrt{\frac{[v^2]}{n(n-1)}} \tag{232}$$

Peters' Formulas:

$$\eta_0=\frac{[v}{n\sqrt{n-1}} \tag{233}$$

$$\epsilon_0=1.2533\frac{[v}{n\sqrt{n-1}} \tag{234}$$

$$r_0=0.8453\frac{[v}{n\sqrt{n-1}} \tag{235}$$

The values of the factors of $[v^2]$ and $[v$ in these formulas are tabulated for various values of n to facilitate computation. Such a table for (231) will be found in Appendix F, Table IV.

142. Example: Precision of the Mean. Let us consider the problem in Art. 28, consisting of 16 observations. Here, $n=16$, $[v=55$, and $[v^2]=305$, the unit being in the fifth place of decimals. The results are as follows:

	ϵ	ϵ_0	r	r_0
Bessel	4.5	1.1	3.0	0.8
Peters	4.5	1.1	3.0	0.7

The range of the observations is 16; one-sixth of this gives 3 as the approximate value of the probable error of a single observation. Using mean square errors, then, the best value from the set of 16 observations would be (from Art. 28),

$$1463.49764\pm0.00001$$

143. Precision of the Weighted Mean. Since the weights are merely relative quantities, as explained in Art. 31, we shall consider them as reduced to integers. The weight of any observation will then be regarded as the number of elemental observations of weight unity of which that observation is the mean. The mean square error, ϵ_1, of an observation of weight w_1, then, will be that of an observation of unit weight, namely, ϵ', divided by $\sqrt{w_1}$, from (229):

$$\epsilon_1 = \frac{\epsilon'}{\sqrt{w_1}}, \qquad \epsilon_2 = \frac{\epsilon'}{\sqrt{w_2}}, \ldots \tag{236}$$

and

$$\epsilon_1\sqrt{w_1} = \epsilon_2\sqrt{w_2} = \epsilon_3\sqrt{w_3} = \quad . \;\; . \;\; . \quad = \epsilon' \tag{237}$$

whence,

$$\frac{\epsilon_1^2}{\epsilon_2^2} = \frac{w_2}{w_1} \tag{238}$$

which states the fundamental principle that the *weights are inversely as the squares of the mean square (or probable) errors.* Also, since the weight of the weighted mean is, by the definition of weights, $[w]$, from (237) we have,

$$\epsilon_0 = \frac{\epsilon'}{\sqrt{[w]}} \tag{239}$$

which corresponds to (229).

To find the expression for ϵ', the mean square error of a single observation of weight unity, we proceed as in the case of equal weights, Art. 140. Beginning with equations (215) and using the first one only, to illustrate the process,

$$\Delta_1 = v_1 + \Delta_0, \quad \Delta_2 = v_2 + \Delta_0, \ldots \tag{215}$$

Squaring,

$$\Delta_1^2 = v_1^2 + 2\Delta_0 v_1 + \Delta_0^2 \tag{240}$$

Multiplying each equation by its weight,

$$w_1\Delta_1^2 = w_1 v_1^2 + 2\Delta_0 w_1 v_1 + w_1\Delta_0^2 \tag{241}$$

. . . .

Since w_1 represents the number of elemental observations of unit weight which make up the first actual observation of weight w_1, it will also be the number of the errors Δ_1, so that $w_1\Delta_1^2$ would be the sum of the squares of these elemental errors; also, by definition, this sum is equal to the number, w_1, times the corresponding mean square error squared, and therefore,

$$w_1\Delta_1^2 = w_1\epsilon_1^2 = \epsilon'^2 \quad \text{from (236)} \tag{242}$$

. . . .

in which ϵ_1 is the mean square error of an observation of weight w_1. Hence,

$$\epsilon'^2 = w_1 v_1^2 + 2\Delta_0 w_1 v_1 + w_1 \Delta_0^2 \tag{243}$$

. . . .

Adding the n equations of this kind, and assuming as in Art. 140 that $\Delta_0 = \epsilon_0$,

$$n\epsilon'^2 = [wv^2] + 2\Delta_0[wv] + [w]\epsilon_0^2 \tag{244}$$

But, from (12), $[wv]=0$,
Therefore,

$$n\epsilon'^2 = [wv^2] + [w]\epsilon_0^2 \tag{245}$$

$$= [wv^2] + \epsilon'^2 \quad \text{by (239)} \tag{246}$$

Whence,

$$\epsilon'^2 = \frac{[wv^2]}{n-1} \tag{247}$$

and the mean square error of an observation of weight unity is,

$$\epsilon' = \sqrt{\frac{[wv^2]}{n-1}} \tag{248}$$

Then from (239),

$$\epsilon_0 = \frac{\epsilon'}{\sqrt{[w]}} = \sqrt{\frac{[wv^2]}{[w](n-1)}} \tag{249}$$

and using the relations stated in (211), we have,

$$r' = 0.6745\sqrt{\frac{[wv^2]}{n-1}} \tag{250}$$

$$r_0 = 0.6745\sqrt{\frac{[wv^2]}{[w](n-1)}} \tag{251}$$

$$\eta' = 0.7979\sqrt{\frac{[wv^2]}{n-1}} \tag{252}$$

$$\eta_0 = 0.7979\sqrt{\frac{[wv^2]}{[w](n-1)}} \tag{253}$$

If the weights are equal, $w=1$, $[w]=n$, and the last six formulas correspond to those of Arts. 140 and 141.

By analogy, the Peters' Formulas for weighted observations may be written. They are,

$$\eta' = \frac{[\sqrt{wv}]}{\sqrt{n(n-1)}} \tag{254}$$

$$\eta_0 = \frac{[\sqrt{wv}]}{\sqrt{[w]n(n-1)}} \tag{255}$$

$$\epsilon' = 1.2533\frac{[\sqrt{wv}]}{\sqrt{n(n-1)}} \tag{256}$$

$$\epsilon_0 = 1.2533\frac{[\sqrt{wv}]}{\sqrt{[w]n(n-1)}} \tag{257}$$

$$r' = 0.8453\frac{[\sqrt{wv}]}{\sqrt{n(n-1)}} \tag{258}$$

$$r_0 = 0.8453\frac{[\sqrt{wv}]}{\sqrt{[w]n(n-1)}} \tag{259}$$

144. Example: Precision of the Weighted Mean. In the problem of Art. 33, $[w]=11$, $n=4$, $[wv^2]=4247$, and the mean square error of the weighted mean is, $\epsilon_0=0.11''$. The complete result is,

$$x_0 = 73° \ 18' \ 42.07'' \pm 0.11''$$

145. Precision of Indirect Observations. The process of finding the mean square errors of the best values of the unknowns from indirect observations is much more involved than in the case of direct observations. Also, the precision is required in comparatively few cases in which engineers are concerned. The method will be outlined, however, without developing the complete theory, for which the reader is referred to the works by Jordan, and Wright and Hayford.

The determination of the precision of the results from indirect observations is divided into two parts, namely, (*a*) the computation of the relative weights of the adjusted values of the unknowns, X, Y, Z, etc., and (*b*) the determination of the mean square error, ϵ', of an observation of weight unity. Then the mean square errors of these unknowns are obtained from the relation (237):

$$\epsilon_x\sqrt{w_x} = \epsilon_y\sqrt{w_y} = \ . \ . \ . \ = \epsilon'$$

or,

$$\epsilon_x^2 w_x = \epsilon_y^2 w_y = \ . \ . \ . \ = \epsilon'^2$$

146. Weights of the Unknowns. There are three methods of determining the weights of the unknowns. We shall use the following one, which utilizes the principle of undetermined coefficients. In the normal equations (22), page 29, to find the weight of X, replace the constant term of the X (i.e., first) equation, $[al]$, by -1, and the other constant terms by zeros. The solution of the equations thus modified will give as the value of X, the reciprocal of its weight, that is, $1/w_x$. Similarly, substituting -1 for $[bl]$ in the second equation, and zeros for the other constant terms, and solving the set of equations for Y, we obtain $1/w_y$. Thus, -1 is substituted for each constant term in succession, the others being replaced by zeros, and the equations are solved, in each case, for the corresponding unknown, the resulting value of which is its $1/w$.

This process is tedious at best, but it can be simplified as follows. It is evident that as the constant terms, only, are altered, the preceding columns of the elimination in the solution of the normal equations will be unchanged. Therefore, referring to the equations (55) page 47, we may add as many columns as there are unknowns, between (l) and (s), designating them as (x_1), (y_2), (z_3), etc., in which to write the new constant terms. These would be included in the check-terms (s') and carried through the elimination the same as other coefficients or constants. Then the weight of each unknown would be determined by substituting back in the proper column until that unknown was determined, and taking the reciprocal of its value. Of course it would be unnecessary to substitute farther in that particular column, as but one weight is obtained from each column. This arrangement of the equations (55) would be,

	x	y	z	(l)	(x_1)	(y_2)	(z_3)	(s')	
(1)	$+6$	-2	$+3$	$+2$	-1	0	0	$+8$	
(2)		$+3$	-1	-3	0	-1	0	-7	(260)
(3)			$+3$	$+1$	0	0	-1	$+2$	

Since the last equation in the elimination will have the quantity -1 as its absolute term in its added column, it follows that the coefficient of the last unknown, in that equation, will always be its own weight. The last added column, (z_3), in the above example, may therefore be omitted.

If the original observations are of unequal weight, the same process is followed, using (21) instead of (22) as the form for the normal equations, and replacing the terms $[wal]$, $[wbl]$, etc., by -1, and zeros, as above.

147. Precision of an Observation of Weight Unity. Let the number of unknowns be represented by m, the number of observations being n, as usual. Then the formulas for the mean square and probable errors of an observation of unit weight, are,

$$\epsilon' = \sqrt{\frac{[wv^2]}{n-m}} \quad \text{or} \quad \sqrt{\frac{[v^2]}{n-m}}$$

$$r' = 0.6745\sqrt{\frac{[wv^2]}{n-m}} \quad \text{or} \quad 0.6745\sqrt{\frac{[v^2]}{n-m}} \tag{261}$$

Lüroth's Formulas are,

$$\epsilon' = 1.2533\frac{[\sqrt{wv}]}{\sqrt{n(n-m)}} \quad \text{or} \quad 1.2533\frac{[v]}{\sqrt{n(n-m)}}$$

$$r' = 0.8453\frac{[\sqrt{wv}]}{\sqrt{n(n-m)}} \quad \text{or} \quad 0.8453\frac{[v]}{\sqrt{n(n-m)}} \tag{262}$$

If there is but one unknown, $m=1$ and these formulas become those of Bessel and Peters for direct observations (Art. 140).

The usual method of determining the residuals is to substitute the adjusted values of the unknowns back into the observation equations and obtain a residual for each equation. However, the sum of the squares of the residuals or of the weighted residuals, that is, $[v^2]$ or $[wv^2]$, may be obtained more easily, in most cases, in the following manner, along with the solution of the normal equations. Form the term at the foot of the diagonal, namely, $[l^2]$ or $[wl^2]$, and perform a corresponding step in the elimination as if there were more terms following it. The resulting sum in the

l-column will then be $[v^2]$ or $[wv^2]$, as the case may be. Also, it may be obtained from the relation

$$[wv^2]=[wal]x+[wbl]y+[wcl]z+\ .\ .\ .\ +[wl^2] \tag{263}$$

and from,

$$[wv^2]=[wvl] \tag{264}$$

which latter requires that the separate v's be known.

148. Example: Precision of Indirect Observations. To illustrate the foregoing articles, the modified observation equations (43), page 38, will be solved to determine the best values of the unknowns and their mean square errors. To the normal equations (44) are added the term $[wl^2]$ and the columns (x_1), (y_2), and (z_3), and the check terms are modified to include all of these.

	x	y	z'	(l)	(x_1)	(y_2)	(z_3)	(s')
I	+3.94	+ 0.38	+2.18	+1.00	−1	0	0	+ 6.50
		+13.56	+0.19	+3.53	0	−1	0	+16.66
			+6.40	+3.09	0	0	−1	+10.86
				+2.66				+10.28
		+13.56	+0.19	+3.53	0	−1	0	+16.66
		− 0.04	−0.21	−0.10	+0.10	0	0	− 0.63
II		+13.52	−0.02	+3.43	+0.10	−1.00	0	+16.03
			+6.40	+3.09	0	0	−1	+10.86
			−1.21	−0.55	+0.55	0	0	− 3.59
			0	+0.01	0	0	0	+ 0.02
III			+5.19	+2.55	+0.55	0	−1.00	+ 7.29
				+2.66				+10.28
				−0.25	+0.25	0	0	− 1.65
				−0.87	−0.03	+0.25	0	− 4.07
				−1.25	−0.27	0	+0.49	− 3.58
IV				+0.29	−0.05	+0.25	+0.49	+ 0.98
			$[wv^2]=$	0.29				

UNKNOWNS

$$z' = \frac{-2.55}{5.19} = -0.491$$

$$y = \frac{-3.43-0.10}{13.52} = \frac{-3.53}{13.52} = -0.261$$

$$x = \frac{-1.00+1.07+0.10}{3.94} = \frac{+0.17}{3.94} = +0.043$$

RESIDUALS

Substituting in the observation equations (43) and determining the v's, we find directly, $[wv^2]=0.28$ and $[wvl]=0.26$, while the evaluation of (263) gives $[wv^2]=0.26$. From the above elimination, the first term of IV is $[wv^2]=0.29$. The average value, is therefore, .027.

WEIGHTS

$$z_3 = \frac{+1}{5.19} \quad \text{and} \quad w_{z'} = 5.2$$

$$z_2 = \frac{0}{5.19} = 0$$

$$y_2 = \frac{+1}{13.52} \quad \text{and} \quad w_y = 13.5$$

$$z_1 = \frac{-0.55}{5.19} = -0.106$$

$$y_1 = \frac{-0.10}{13.52} = -0.007$$

$$x_1 = \frac{+1.00+0.23}{3.94} = \frac{1}{3.2} \quad \text{and} \quad w_x = 3.2$$

MEAN SQUARE ERRORS

Average value of $[wv^2]=0.27$; $n=9$; $m=3$.

$$\epsilon' = \sqrt{\frac{0.27}{6}} = \sqrt{0.0475} = 0.22$$

$$\epsilon_x = \frac{\epsilon'}{\sqrt{w_x}} = \frac{0.22}{1.8} = 0.12$$

$$\epsilon_y = \frac{0.22}{3.7} = 0.06$$

$$\epsilon_{z'} = \frac{0.22}{2.3} = 0.10$$

Results

$$x = +0.043 \pm 0.12$$
$$y = -0.261 \pm 0.06$$
$$z' = -0.491 \pm 0.10$$

149. Precision of Conditioned Observations. In general, it is necessary, as in the preceding case, to determine the precision of an observation of weight unity and also the weight of each unknown, from which the precision of the unknowns is obtained from the usual relation that the mean square errors are inversely proportional to the square roots of the weights, that is, from,

$$\epsilon_x^2 w_x = \epsilon_y^2 w_y = \quad . \quad . \quad . \quad = \epsilon'^2$$

If the conditioned observations are adjusted as *indirect observations* by the method stated in Art. 81, the precision of those unknowns which are involved in the normal equations may be determined by the methods just explained in Arts. 145 to 148. Then by a second solution, eliminating a different set of unknowns, the normal equations may be made to involve those which were not included in the previous set, and their precision may be found in the same manner. Obviously, this is a tedious method except in cases of a few observations.

Since the number of unknowns which may thus be made independent is the total number, m, minus the number of conditions, m', the formula for the mean square error of a single observation of weight unity may be derived directly from (261) and (262) by substituting for m, $m-m'$. Thus,

$$\epsilon' = \sqrt{\frac{[wv^2]}{n-m+m'}} \quad \text{or} \quad \sqrt{\frac{[v^2]}{n-m+m'}}$$

$$r' = 0.6745\sqrt{\frac{[wv^2]}{n-m+m'}} \quad \text{or} \quad 0.6745\sqrt{\frac{[v^2]}{n-m+m'}} \qquad (265)$$

But in most of the cases with which we are concerned each unknown is directly observed so that $n = m$, when the above formulas become,

$$\epsilon' = \sqrt{\frac{[wv^2]}{m'}} \quad \text{or} \quad \sqrt{\frac{[v^2]}{m'}}$$

$$r' = 0.6745\sqrt{\frac{[wv^2]}{m'}} \quad \text{or} \quad 0.6745\sqrt{\frac{[v^2]}{m'}} \qquad (266)$$

in which m' represents the number of conditions. Lüroth's Formulas (262) may be similarly modified by substituting $m-m'$ for m; and when $n=m$, the denominators become $\sqrt{nm'}$.

The residuals, v, are the corrections, v, to the observations, as determined in the adjustment (Art. 73). As a check upon the direct computation of $[wv^2]$, however, we may use the formula,

$$[wv^2] = -Aq_1 - Bq_2 - Cq_3 \quad . \quad . \quad . \tag{267}$$

A, B, C, . . . being the correlates, and q_1, q_2, q_3, . . . being the absolute terms of the reduced condition equations (59) or the normal equations (64). Or, in the solution of the normal equations, a step may be taken similar to the one described in Art. 147 for indirect observations. Here, however, zero is written for the last term in the constant (q) column. The elimination process is continued to include this term, and the resulting sum will be $-[wv^2]$.

The *method of correlates*, however, will generally be used in the adjustment. The weights of the adjusted values are not determined directly, in this case, but the weight of a function of these values is determined, and this function may be merely unity times one of them. Examples of the functions of the adjusted values for which the precision may be desired are: A side of a triangle or an unobserved line in a system of triangulation, when computed from the adjusted angles; and a computed difference of elevation in a level net, determined from adjusted values of observed differences. The function must not involve more of the unknowns than can be made independent by elimination with the conditions, that is, not more than $m+m'$. The method is as follows:[1]

Since any function can be reduced to the linear form, this one will be assumed to have that form,

$$F = f_0 + f_1V_1 + f_2V_2 + f_3V_3 + \quad . \quad . \quad . \quad + f_nV_n \tag{268}$$

in which V_1, V_2, V_3, . . . are the adjusted values of the unknowns (Art. 72), namely, $V_1 = M_1 + v_1$, etc. If any of the terms in (268) are missing from the desired function, give to the corresponding

[1] See Jordan, Handbuch der Vermessungskunde, Bd. I, par. 45, or Wright's Adjustment of Observations, page 229.

coefficients, f, the value zero. Now, referring to the condition equations (56) or (59) for the notation, and representing the original weights of the observations by $w_1, w_2, w_3, \ldots$, we form the terms, $m'+1$ in number,

$$\left[\frac{af}{w}\right], \left[\frac{bf}{w}\right], \left[\frac{cf}{w}\right], \ldots \left[\frac{ff}{w}\right] \tag{269}$$

Writing these terms in order in an additional column, between the constant and check in the normal equations, or in place of the constant column if the equations have already been solved, the elimination process is carried out for this column and its own new check column, including the last step, for the term $[ff/w]$. The final sum for the last term will be the reciprocal of the weight of the function,[1] and from this and the mean square error of a single observation of unit weight, the mean square error of the function is obtained by (237),

$$\epsilon_F = \frac{\epsilon}{\sqrt{w_F}} \tag{270}$$

It will be seen that even though the weights of the original observations were equal, those of the adjusted values may be unequal. But if the original weights of certain of the observations are equal, and also $a, b, c, \ldots$ are the same for all of these observations, then the weights of the corresponding adjusted values will be equal, since the f's in these cases are unity. This will appear in the following examples.

150. Examples: Differences of Elevation. (a) As an illustration of the method, let us apply it to the problem of Art. 77, Adjustment of Levels, and determine the mean square error of the difference of elevation of the benchmarks A and D. The function is, therefore,

$$F = V_6 + V_8 \tag{271}$$

[1] Expressed in the Gaussian form, this is,

$$\frac{1}{w_F} = \left[\frac{ff}{w}\right] - \frac{\left[\frac{af}{w}\right]^2}{\left[\frac{a^2}{w}\right]} - \frac{\left[\frac{bf}{w}\cdot 1\right]^2}{\left[\frac{b^2}{w}\cdot 1\right]} - \frac{\left[\frac{cf}{w}\cdot 2\right]^2}{\left[\frac{c^2}{w}\cdot 2\right]} - \ldots$$

and the reciprocals of the weights of the observations are, $1/w_6=2$, and $1/w_8=1$. Then $f_6=+1$, and $f_8=+1$, the other f's being zero. Referring to the condition equations (73), we find

$$\begin{array}{ll} a_6=-1 & a_8=0 \\ b_6=-1 & b_8=0 \\ c_6=0 & c_8=-1 \end{array}$$

Thus we obtain the terms,

$$\left[\frac{af}{w}\right]=-2,\quad \left[\frac{bf}{w}\right]=-2,\quad \left[\frac{cf}{w}\right]=-1,\quad \left[\frac{ff}{w}\right]=+3 \qquad (272)$$

Utilizing the solution of the normal equations already made in Art. 77, we write the f-terms in a new column and form a new check column. To illustrate the second method of obtaining $[wv^2]$, stated near the middle of page 173, the constant column will be included, with the addition of the zero for the last term. The solution, then, is as follows:

	A	B	C	(f)	Check	Const.	
1	+12	+6	−4	−2	+12	+0.09	
2		+7	0	−2	+11	+ .08	
3			+5	−1	0	− .08	
4				+3	− 2	0	
2		+7	0	−2	+11	+ .08	
5		−3	+2	+1	− 6	− .045	(I)×(−6/12)
II		+4	+2	−1	+ 5	+ .035	(2)+(5)
3			+5	−1	0	− .08	
6			−1.33	−0.67	+ 4	+ .0300	(I)×(+4/12)
7			−1	+0.50	−2.50	− .0175	(II)×(−2/4)
III			+2.67	−1.17	+1.50	− .0675	(3)+(6)+(7)
4				+3	−2	0	
8				−0.33	+2	− .0007	(I)×(+2/12)
9				−0.25	+1.25	− .0003	(II)×(+1/4)
10				−0.51	+0.66	− .0017	(III)×(−1.17/2.67)
IV				+1.91	+1.91	−0.0027	

Therefore,

$$\frac{1}{w_F}=1.91 \tag{273}$$

and

$$-[wv^2]=-0.0027$$

From (76), we evaluate (267) and obtain,

$$[wv^2]=-0.0012+0.0018+0.0021=+0.0027$$

which agrees with the above and with the value determined directly from the table of corrections, page 69.

Then,

$$\epsilon'=\sqrt{\frac{[wv^2]}{m'}}=\sqrt{\frac{0.0027}{3}}=0.03 \tag{274}$$

and from (272) and (273) we have,

$$\epsilon_F=\frac{\epsilon'}{\sqrt{w}}=\frac{0.03}{\sqrt{1.91}}=0.02 \tag{275}$$

so that the best value of the difference of elevation from A to D, from the given data, is, with its mean square error,

$$(F)=-6.36\pm0.02 \tag{276}$$

(*b*) As a variation of the above, let us determine the precision of the adjusted difference of elevation, $A-F$. The function is,

$$F'=V_6 \tag{277}$$

and from the data above,

$$1/w_6=2,\ f_6=+1,\ a_6=-1,\ b_6=-1,\ \text{and}\ c_6=0.$$

so that

$$\left[\frac{af}{w}\right]=-2,\quad \left[\frac{bf}{w}\right]=-2,\quad \left[\frac{cf}{w}\right]=0,\quad \left[\frac{ff}{w}\right]=+2.$$

The only changes in the solution, therefore, are in the last two of the f-terms. The result is

$$\frac{1}{w_{F'}}=1.41 \tag{278}$$

whence,

$$\epsilon_{F'}=\frac{\epsilon'}{\sqrt{1.41}}=\frac{0.03}{1.2}=0.02 \tag{279}$$

In view of the statements at the close of Art. 149, it is evident from an inspection of the tabulated condition equations (73) that the weight and mean square error of V_2 will be the same as those of V_6, since these two columns in the table are alike.

151. Precision of Computed Quantities. As a result of the adjustment of observations, the adopted values of the unknowns are likely to be used in the computation of other quantities which may be expressed as functions of the unknowns. Having investigated the precision of the unknowns, it may be desired to ascertain the effect which the uncertainties in these values would have upon the quantities computed from them.

For example, suppose the diameter of a cylindrical bar of steel is measured with micrometer calipers at various points, from which the mean diameter and its probable error are obtained; the cross-sectional area computed from this mean diameter would have a resulting uncertainty. Also, if the bar were tested in a tension machine, the breaking stress per square inch would be uncertain to a corresponding degree as a result of the uncertainty in the measured diameter and computed area.

Again, suppose one side and the adjacent angles of a triangle have been measured independently, resulting in an adopted mean and a mean square error for each. If another side be computed from these data, it will have an uncertainty due to the discrepancies among the original measures of the given side and angles, that is, to the uncertainties of the given means.

It must be emphasized that the determination of the best values of the computed quantities is not involved in this question. Having adjusted the observations, the resulting adopted values are the best ones, as far as our knowledge goes, and quantities computed from them are also the best we can determine *from the given data.* We are now concerned only with the precision of the computed quantities, not with the determination of the quantities themselves.

Our problem is to determine the mean square (or probable) error of a *function* of independent, adjusted quantities of which the mean square (or probable) errors are given. It will be convenient to assume that each of these given, adjusted values is the mean of a large number of observations, and that the corresponding indices of precision were determined by the formulas of Art. 141, although, of course, they might result from indirect observations.

The errors in a linear function of independently observed quantities occur in accordance with the same Law of Error as those of the quantities themselves.[1] Thus, the errors in the mean of a set of observations occur in accordance with the usual Law of Error. Such means, therefore, may be treated as original observations, as far as the occurrence of errors goes, as long as they do not involve the same original observations, in which case they would no longer be independent.[2]

This subject is usually called the *Propagation of Error.* We shall consider it as divided into two parts,—the *simple* influence of errors of one kind or character, and the *compound* effects of errors of different kinds or resulting from different causes.

152. Simple Propagation of Error. Before attacking the general case, a few special forms of functions will be considered in order to illustrate the process of reasoning. Let F represent the function of the independent, adjusted quantities, x, y, . . . whose mean square errors are ϵ_x, ϵ_y, . . . Let the original observations of x be represented by M_1, M_2, . . ., those of y by M'_1, M'_2, . . ., etc., and let the true errors of these observations be represented respectively by Δ_1, Δ_2, . . ., Δ'_1, Δ'_2, . . ., etc. We may assume an equal number of observations for each quantity, for simplicity.

(a) Consider first the sum or difference of two quantities. Then,

$$F = x \pm y \tag{280}$$

Taking the separate observations in pairs, the first of x with the first of y, the second of x with the second of y, etc., each pair gives a value of F, say F_1, F_2, . . . Thus,

$$\begin{aligned} F_1 &= M_1 \pm M'_1 \\ F_2 &= M_2 \pm M'_2 \\ &\cdots\cdots\cdots \end{aligned} \tag{281}$$

[1] For proof of this, see Wright and Hayford, Adjustment of Observations, Art. 13.

[2] See Chauvenet par. 23, for treatment of the case of a function of functions.

Now, if we add to each M its true error, Δ, the resulting value of F must be corrected by its corresponding error, and,

$$\begin{aligned} F_1+\Delta_{F_1} &= (M_1+\Delta_1)\pm(M'_1+\Delta'_1) \\ F_2+\Delta_{F_2} &= (M_2+\Delta_2)\pm(M'_2+\Delta'_2) \\ &\ldots\ldots\ldots\ldots \end{aligned} \tag{282}$$

Subtracting (281) from (282), one by one, we have,

$$\begin{aligned} \Delta_{F_1} &= \Delta_1\pm\Delta'_1 \\ \Delta_{F_2} &= \Delta_2\pm\Delta'_2 \\ &\ldots\ldots \end{aligned} \tag{283}$$

Squaring each equation, adding, and dividing by their number, n,

$$\frac{[\Delta_F^2]}{n}=\frac{[\Delta^2]}{n}\pm\frac{2[\Delta\Delta']}{n}+\frac{[\Delta'^2]}{n} \tag{284}$$

But in a large number of observations, the positive and negative errors occur with equal frequency, so that the sum of the products, $[\Delta\Delta']$, would approximate to zero,—certainly so in comparison to $[\Delta^2]+[\Delta'^2]$, so that,

$$\frac{[\Delta_F^2]}{n}=\frac{[\Delta^2]}{n}+\frac{[\Delta'^2]}{n} \tag{285}$$

or,

$$\epsilon_F^2=\epsilon_x^2+\epsilon_y^2 \tag{286}$$

Obviously, the above process would apply likewise to a similar function consisting of any number of quantities connected by plus and minus signs, so that for

$$F=x\pm y\pm z\pm \ \ldots$$

we can write,

$$\epsilon_F^2=\epsilon_x^2+\epsilon_y^2+\epsilon_z^2+ \ \ldots \tag{287}$$

and

$$\epsilon_F=\sqrt{\epsilon_x^2+\epsilon_y^2+\epsilon_z^2+ \ \ldots} \tag{288}$$

From the constant ratio of the probable error to the mean square error, it follows that

$$r_F^2=r_x^2+r_y^2+r_z^2+ \ \ldots \tag{289}$$

This principle is very important and often used. Note that the signs in (287) and (289) are all positive. The uncertainty in the sum of two or more quantities is therefore the same as in their difference.

(*b*) In the next case, let $F=ax\pm by\pm \ \ldots$ (290)

in which x and y are adjusted values from observations, and a and b are known constants. As in (282), we may write,

$$
\begin{aligned}
F_1+\Delta_{F_1}&=a(M_1+\Delta_1)\pm b(M'_1+\Delta'_1)+\ \ldots \\
F_2+\Delta_{F_2}&=a(M_2+\Delta_2)\pm b(M'_2+\Delta'_2)+\ \ldots \\
&\ldots\ldots\ldots\ldots
\end{aligned}
\tag{291}
$$

whence, as in (283),

$$
\begin{aligned}
\Delta_{F_1}&=a\Delta_1\pm b\Delta'_1+\ \ldots \\
\Delta_{F_2}&=a\Delta_2\pm b\Delta'_2+\ \ldots \\
&\ldots\ldots\ldots\ldots
\end{aligned}
\tag{292}
$$

Squaring, adding, dividing by n, and omitting the products as before, we have,

$$\frac{[\Delta_F{}^2]}{n}=a^2\frac{[\Delta^2]}{n}+b^2\frac{[\Delta'^2]}{n}+\ \ldots \tag{293}$$

That is,

$$\epsilon_F{}^2=a^2\epsilon_x{}^2+b^2\epsilon_y{}^2+\ \ldots \tag{294}$$

or,

$$\epsilon_F=\sqrt{a^2\epsilon_x{}^2+b^2\epsilon_y{}^2+\ \ldots} \tag{295}$$

(*c*) Now we shall consider the general case in which F is *any* function of the quantities, x, y, z, etc.,

$$F=f(x,\ y,\ z,\ \ldots) \tag{296}$$

Since x, y, z, . . . are adjusted values, they may be assumed to be nearly correct, so that their errors are very small; let us represent them by differentials. Then, if Δ_F be the true error of F, we have,

$$F+\Delta_F=f(x+dx,\ y+dy,\ z+dz,\ \ldots) \tag{297}$$

Expanding this function by Taylor's Theorem, and omitting terms which involve squares, products, and higher powers of the differentials, we obtain,

$$F+\Delta_F=f(x,\ y,\ z,\ \ldots)+\frac{\partial f}{\partial x}dx+\frac{\partial f}{\partial y}dy+\frac{\partial f}{\partial z}dz+\ \ldots \tag{298}$$

whence, subtracting (296),

$$\Delta_F=\frac{\partial f}{\partial x}dx+\frac{\partial f}{\partial y}dy+\frac{\partial f}{\partial z}dz+\ \ldots \tag{299}$$

This equation has the same linear form as (292), so that from (294) we have directly,

$$\epsilon_F{}^2=\left(\frac{\partial f}{\partial x}\right)^2\epsilon_x{}^2+\left(\frac{\partial f}{\partial y}\right)^2\epsilon_y{}^2+\left(\frac{\partial f}{\partial z}\right)^2\epsilon_z{}^2\ \ldots \tag{300}$$

in which the partial derivatives of the function correspond to the constants, a, b, c, etc., of (294).

Thus we may state the general *Rule:* Express the given function in literal form. Differentiate it partially with respect to each quantity for which the mean square error is given. Substitute in these derivatives the given quantities (without reference to their mean square errors, of course). Substitute in (300) and obtain ϵ_F.

153. Example: Precision of the Mean. We shall now apply the foregoing principles to determine the mean square error, ϵ_0, of the mean of n observations, when the mean square error of a single observation is ϵ.

The expression for the mean is,

$$F = x_0 = \frac{M_1}{n} + \frac{M_2}{n} + \quad . \quad . \quad . \quad + \frac{M_n}{n} \tag{301}$$

where M_1, M_2, etc., represent independent direct measures or observations of the unknown quantity. This function has the form of (290) and $a = b = c = \quad . \quad . \quad . \quad = 1/n$. From (294), therefore, we have

$$\epsilon_0^2 = \frac{1}{n^2}\epsilon^2 + \frac{1}{n^2}\epsilon^2 + \frac{1}{n^2}\epsilon^2 + \quad . \quad . \quad \text{to } n \text{ terms} \tag{302}$$

that is,

$$\epsilon_0^2 = n\left(\frac{\epsilon^2}{n^2}\right) = \frac{\epsilon^2}{n} \tag{303}$$

or,

$$\epsilon_0 = \frac{\epsilon}{\sqrt{n}} \tag{304}$$

which states the very important principle that the precision of the mean varies directly as the square root of the number of observations. To double the precision,—that is, to reduce the mean square error of the mean to one-half its size,—it is necessary to have four times as many observations.

154. Compound Propagation of Error. The uncertainty in a computed quantity may result from several sources which are not of the same nature, and it may be impossible to state the quantity as a single function of all these sources of error. For example, the measurement of a line with a steel tape involves the uncertainty in the length of the tape itself and also the errors in the process of measurement. We cannot express the length of the line as a function of the length of the tape and the " process of measurement!"

From (283) to (287), we can state the principle that when the error in the computed quantity is the algebraic sum of independent errors from different sources, the total mean square error of the computed quantity will be the square root of the sum of the squares of the separate mean square errors of that quantity due to the various causes.

In any given case, therefore, we determine the mean square error of the computed quantity or function resulting from each source, separately, by the methods of Art. 152, and then take the square root of the sum of their squares as the total mean square error.

We shall now illustrate this subject by a series of typical examples.

155. Examples: Propagation of Error. (1) The following measures of the diameter of a cylindrical test-piece of metal were made by means of micrometer calipers. The piece was then broken in a testing-machine at a load of 20,000 lbs. Find the unit breaking stress and its mean square error due to the uncertainty in the measured diameter, D.

Inches	v	v^2	Inches	v	v^2	Inches	v	v^2
0.6252	2	4	0.6251	1	1	0.6248	2	4
46	4	16	52	2	4	45	5	25
42	8	64	54	4	16	52	2	4
48	2	4	55	5	25	56	6	36

$$\frac{121}{12}=10;\quad \text{Mean}=0.6240+0.0010=0.6250=D$$

$$[v^2]=203,\quad \epsilon_0{}^2=\frac{203}{132}=1.54,\quad \epsilon_0=1.24$$

Diameter $(D)=0.6250\pm0.00012$. The function is,

$$\text{Unit breaking stress}=\frac{20,000}{\text{Area of section}}$$

$$=\frac{20,000}{\frac{\pi D^2}{4}}$$

$$=\frac{80,000}{3.14}\left(\frac{1}{D^2}\right) \tag{305}$$

$$=65,300 \text{ lbs. per sq. in.}$$

Differentiating (305) with respect to D,

$$\frac{\partial f}{\partial D}=\frac{80,000}{3.14}\left(\frac{-2}{D^3}\right)=\frac{-160,000}{3.14\times 0.625^3}$$

$$=\frac{-160,000}{0.766}=-208,900$$

From (300),

$$\epsilon_F=208,900\times 0.00012=25.$$

Thus, the unit breaking stress $=65,300\pm 25$ lbs. per square inch.

The uncertainty due to the variation among the measures of the diameter is therefore negligible when it is remembered that the breaking load is seldom required within a range of a hundred pounds.

(2) The length of a 50-meter tape is determined by comparison with a 5-meter standard bar which is surrounded by chipped ice to control its temperature. The length of the bar, as determined from its standardization, is $B=5.000060\pm 0.000006$ meters. The following measures are made of the difference between the length of the tape and ten lengths of the bar, the former being the longer. It is required to find the length of the tape and its mean square error due to the uncertainty in the length of the bar and to the errors of measurement. The temperature is assumed constant. The unit is in the sixth place of decimals, that is, a micron.

Interval (K) Obs.	r	r^2
0.002533	16	256
666	117	13689
529	20	400
444	105	11025
461	88	7744
553	4	16
638	89	7921
567	18	324
Mean, 2549		41375

$$\epsilon_0^2=\frac{41375}{56}=739. \qquad \text{Interval } (K)=2549\pm 27$$

The function is,

$$\text{Length of tape } (L) = 10\ B + K \tag{306}$$

$$= 50.000600 + 2549$$

$$= 50.003149 \text{ m.}$$

This function corresponds to (290), so that, from (294),

$$\epsilon_L{}^2 = 100 \times 6^2 + 739 = 3600 + 739 = 4339.$$

$$\epsilon_L = 66$$

Therefore, Length of tape $(L) = 50.003149 \pm 0.000066$ m.

An important principle is illustrated here. The larger source of error in the length of the tape is that due to the error in the length of the bar, amounting to ten times as much as the other. It would be useless, therefore, to increase the above number of observations with the idea of increasing the precision in the length of the tape, since this part of the total error is almost negligible. On the other hand, the above set of observations might be diminished considerably without seriously affecting the result. For example, suppose there were but one-half as many observations, namely, 4. Dividing the number of observations by 2 increases the square of the mean square error twofold. Thus, we should have $\epsilon_0{}^2 = 1478$, and $\epsilon_L{}^2 = 3600 + 1478 = 5078$. Then, $\epsilon_L = 71$, which is very little larger than 66. It must be remembered however, that the number of observations should be sufficiently large to justify the assumption that errors of observation follow the Law of Error.

(3)[1] A comparison of the two following cases will be instructive. (*a*) A line 400 feet long is measured with a 100-foot tape of which the mean square error is 0.004 foot. The resulting mean square error in the length of the line will be 0.016 foot, since $L = 4T$.

(*b*) The same line is divided into four 100-foot sections and each section is measured with a different 100-foot tape of which the mean square error is 0.004 in each case. The resulting mean

[1] Adapted from Crandall's Geodesy and Least Squares.

square error in the length of the line will be $\sqrt{4(0.004)}=0.008$ foot, since the function is, $L=T_1+T_2+T_3+T_4$.

In the first case (a), whatever the *true* error in the tape may be, it is constant and its effect is cumulative. In (b), on the other hand, the *actual* errors in the different tapes are not the same even though their mean square errors happen to be equal, and in consequence they are likely to be both positive and negative so as to neutralize to some extent. Therefore, the resulting error in the length of the line would be smaller than in the former case. It is important that this principle be well understood.

(4) Let it be required to compute the length and mean square error of the side, b, of the triangle, A-B-C, from the side, a, and the angles, A and B, given with their mean square errors as follows:

$$a=4268.344\pm0.008 \text{ meter},$$

$$A=56° \ 37' \ 42.4''\pm0.6''$$

$$B=70° \ 26' \ 54.3''\pm0.3''$$

The function is,

$$b=\frac{a \sin B}{\sin A} \tag{307}$$

from which we obtain, using the above data,

$$b=4816.349 \text{ m}.$$

Differentiating (307) with respect to a, A, and B, in succession, and reducing by means of (307),

$$\frac{\partial b}{\partial a}=\frac{\sin B}{\sin A}=\frac{b}{a}=1.128$$

$$\frac{\partial b}{\partial A}=\frac{-a \sin B \cos A}{\sin^2 A}=-b \cot A=-3172$$

$$\frac{\partial b}{\partial B}=\frac{a \cos B}{\sin A}=b \cot B=1710$$

Substituting in (300), and noting that it is necessary to multiply ϵ_A and ϵ_B by $\sin 1''$ ($=0.000005$) in order to reduce them to abstract quantities so that each term may be expressed in the unit of length, we have,

$$\epsilon_b^2 = \left(\frac{b}{a}\right)^2 \epsilon_a^2 + (b \cot A)^2(\epsilon_A \sin 1'')^2 + (b \cot B)^2(\epsilon_B \sin 1'')^2$$

$$= (1.128 \times 0.008)^2 + (3172 \times 0.6 \times 0.000005)^2$$
$$+ (1710 \times 0.3 \times 0.000005)^2$$
$$= (0.0090)^2 + (0.0095)^2 + (0.0026)^2$$
$$= 0.00017801$$

and

$$\epsilon_b = 0.013$$

whence,

$$b = 4816.349 \pm 0.013 \text{ meters.}$$

(5) Find the mean square error in a single measurement of an angle, direct and reversed, with a direction theodolite having three microscopes. Each reading consists of the mean of the three microscope readings corresponding to a pointing upon one object, and a measure of the angle is the difference between the readings upon the two objects limiting the angle. This process is repeated in the reversed position of the instrument and the mean is taken. Suppose the mean square error of a pointing of the telescope upon an object to be, $\epsilon_p = 0.04''$; that of a reading of a microscope to be, $\epsilon_r = 0.06''$; and that of a graduation-mark on the circle to be, $\epsilon_j = 0.03''$. The error in each microscope reading will be the algebraic sum of the error of setting and reading the microscope itself and that of the graduation, so that the mean square error due to both causes will be $\sqrt{(\epsilon_r^2 + \epsilon_j^2)}$. Then the mean square error of the mean of the readings of the three microscopes will be

$$\epsilon_R = \frac{\sqrt{(\epsilon_r^2 + \epsilon_j^2)}}{\sqrt{3}} = \frac{\sqrt{(0.0045)}}{\sqrt{3}} = \sqrt{0.0015}$$

The error in a reading upon one object will be made up of the errors due to all three causes, that is, to the above combined error and the error of a single pointing, or,

$$\epsilon_0 = \sqrt{(\epsilon_R^2 + \epsilon_p^2)} = \sqrt{(0.0015 + 0.0016)} = \sqrt{0.0031}$$

Finally, the mean square error of the difference of the readings on the two objects, which is that of the direct measurement of the angle, will be,

$$\sqrt{(\epsilon_0^2+\epsilon_0^2)}=\sqrt{0.0062}$$

and that of the mean of the direct and reversed results will be

$$\epsilon_A=\frac{\sqrt{0.0062}}{\sqrt{2}}=\sqrt{0.0031}=0.056''$$

(6) A line 1000 feet long is measured eight times with a 100-foot tape, and the mean square error of the mean of the eight measures is found to be 0.004 foot. If the mean square error of the length of the tape (resulting from its standardization) is 0.001, what is the mean square error of the line, due both to errors of manipulation and error in the tape length?

The mean square error of the line due to the tape error is $10\times 0.001=0.010$. Since the total error is the algebraic sum of both kinds, the mean square error due to both causes will be the square root of the sum of the squares of the separate mean square errors, that is,

$$\epsilon_L=\sqrt{(0.004^2+0.010^2)}=0.011$$

COMBINATION OF COMPUTED QUANTITIES

156. Weights from Mean Square or Probable Errors. In Art. 143, it was demonstrated that weights are inversely as the squares of the corresponding mean square or probable errors. Thus it is possible to combine the results computed from different observations of a certain quantity, using them as weighted observations, when the mean square errors of these results are known so that their relative weights may be determined. For example, a certain angle in a triangulation may have been measured several times, with a resulting mean and mean square error. Subsequently, in another season, perhaps, another series of measures of the same angle may be made, giving a different result and mean square error. By giving to each result a weight equal to the reciprocal of the square of its mean square error, the weighted mean of the two results may be taken as the best value of the angle from all of the available data.

157. Limitations. It is obvious that this method assumes that all of the original observations in the various groups are of the same character, so that if they were known their mean could reasonably be taken. The conditions under which they were made should be similar, and especially is it assumed that constant or systematic errors affect all of them in the same way.

On the other hand, it is seldom that these conditions are fulfilled with any great degree of certainty. Frequently, nothing at all is known about the observational methods or circumstances, except what is indicated by the mean square errors as to the consistency of the original observations. Even in such a case, however, it is probable that the weighted mean will be as good as, or better than, any of the component results, so that the method should not be discarded without careful consideration.

Of especial importance in this connection, is the case in which the observations resulting in one of the given values are known to be of much greater precision than those which resulted in the other value, without regard to their respective mean square errors. For example, an angle might be measured with a direction theodolite reading to a single second, and again by means of a transit reading to half-minutes. Here, the judgment of the computer may determine what weight, if any, shall be given to the transit result in comparison with the other, in spite of the mean square errors, provided, of course, the number of observations made with the theodolite is sufficient to reduce the effect of the accidental errors. Should the two results be close together, however, the weights given by the mean square errors may still be satisfactory.

When the results being compared are separated by a considerable interval in comparison with the given mean square errors, the presence of systematic error may be indicated and should be investigated. If the difference is not too great to be a reasonable accidental error of observation, it may be considered safe to accept the weights given by the mean square errors. But if the difference is too large to be thus considered, and the mean square errors are much smaller, there may be no reason for believing one of the values to be nearer the truth than the other, so that the arithmetic mean of them may be adopted as the best value. Here, again, the judgment of the computer must determine the method of adjustment.[1]

158. Example: Weighted Mean of Computed Quantities. Three independent series of observations give the following results for the value of an angle; what is the best or most probable value of the angle from these data?

Means, (x_0)	ϵ	ϵ^2	$w=1/\epsilon^2$	w	wx_0
72° 47′ 43.18″	±0.06″	36	0.028	14	44.52
44.01	.10	100	.010	5	20.05
43.74	.08	64	.016	8	29.92
				27	)94.49
					3.50

Weighted mean, 72° 47′ 43.50″

159. Precision of the Adjusted Value. Although the problem has the nature of the determination of the weighted mean, the precision of the resulting value should not be computed as in the case of the weighted mean owing to the small number of given quantities, in general, and the fact that their individual mean square errors or probable errors are given. On the contrary, the method of propagation of error should be used. The two methods will not usually give the same result. It is to be expected that the final value will be better than the separate given values, and it should generally have a smaller mean square error although this will not always be the case.

160. Example: Precision of the Adjusted Value. Applying this process to the above example of Art. 158, we have the function,

$$X=\frac{1}{27}\left(14x_1+5x_2+8x_3\right)$$

[1] See Johnson's Theory of Errors, Chap. VII, on this subject.

whence, from (300),

$$\epsilon_X{}^2 = \frac{1}{27^2}[(14\times .06)^2+(5\times .10)^2+(8\times .08)^2]$$

$$= \frac{1.3652}{27^2} = \frac{1.17^2}{27^2}$$

$$\epsilon_X = \frac{1.17}{27} = 0.04$$

so that the adjusted value is, 72° 47′ 43.50″ ± 0.04″.

CHAPTER IX

CONCLUSION

161. Rejection of Observations. It is generally conceded that an observer has the right to reject any observation, at the time of making it, if he has reason to believe that he made a mistake in his setting or reading, or if the conditions were temporarily so unfavorable as to indicate that the result was quite unreliable. His attention may be drawn to the questionable observation merely by its being discordant among the others of the series; or he may question the observation as he makes it and mentally decide to reject it if it proves to be very discordant. His power is absolute but he is expected to exercise it with good judgment and strict impartiality.

On the other hand, when the observations have been approved by the observer and are turned over to the computer, or when sufficient time has elapsed that the observer ceases to recall the particular conditions under which each of the observations was made, then the record must be regarded as inviolable, and must not be changed without good reason,[1] and this reason must be evident from the records themselves.

If the observer has noted the unfavorable conditions and has not indicated a resulting smaller weight for the corresponding observation, the computer may feel justified in assigning such a weight if the observation is clearly discordant. However, if this is necessary, it should have been done by the observer in the field, and the computer may wisely refrain from thus interfering with the record unless with the consent of the observer himself, on the ground that this would have been his action in the field.

[1] It is a rigid rule that an original record should never be erased or obscured. Changes should be so made as to show clearly that they are changes, with date and initials of the computer, and so as to leave the original data legible. Generally, the original will be in pencil, and notes and computations will be in ink. Red ink may well be used for annotations.

The assignment of weights to various observations is closely associated with the question of the rejection of observations, since a weight of zero is equivalent to rejection, and a diminished weight means a partial rejection.

162. Criteria for Rejection of Observations. While the author is of the opinion that weighting and rejection should be based upon judgment rather than mere discrepancies among the observations, many writers and experienced computers have advocated the rejection of all observations which deviate more than a certain amount from the mean of the set. The mathematical basis for determining this maximum deviation is known as a *Criterion for the Rejection of Observations.* Several of these methods have been devised,[1] but the following has the merit of simplicity.

It being assumed that the observations conform to the Law of Error, the number of errors, or residuals, greater than a certain size, to be expected in the given set, will be found by using Table III, Appendix F, as stated in Art. 175. The table shows that the probability of an error less than four times the probable error of a single observation is 0.99; that is, 99 out of 100 residuals should be less than that amount and only one out of 100 should be greater. Therefore, if a greater residual occurs in a set of, say, 20 to 30 observations, it might be rejected as indicating a mistake. Having computed the probable error, r, of a single observation, for the given set, any individual observation whose residual from the mean is greater than or equal to $4r$, would be rejected, according to this assumption.

Evidently, the adoption of a certain criterion is a matter of estimation and preference. The above value, $4r$, would be considered conservative by many computers who believe in any kind of a numerical criterion; $3r$ is sometimes used. Even the novice will immediately suggest that the unusually large error might *happen* to occur in the small series of observations. If but one very large residual occurs in the set, there may be more reason for rejecting it than if it be accompanied by a correspondingly large one of the opposite sign, since the pair would neutralize each other, to some extent, in the mean.

[1] See Chauvenet, Practical and Spherical Astronomy.

163. Methods of Observing. One of the most important uses of the Method of Least Squares lies in the investigation of methods of observing, with the idea of avoiding or eliminating the effects of constant or systematic influences, of segregating the sources of error which produce the greatest effect so that these effects may be diminished, and of reducing the cost of securing the desired degree of precision.[1]

In Art. 154 it was shown how various sources of error combined to affect the result; therefore, in arranging the observations, special attention should be given to decreasing the errors which have the greatest effect, since the final precision is dependent but little upon the small errors. In reducing the errors from a certain source, the design of the instrument and its support may require study as well as the method of using it. Very important improvements in instruments have resulted from the careful study of the occurrence of the errors of observation.[2]

Constant and systematic errors may be due to the conditions under which the observing is carried on. When such is the case, it is desirable to so arrange the observing program that these conditions will vary during the observations through a complete cycle of changes, as far as practicable, in order that their effects may neutralize one another, at least partially.

Finally, the matter of cost must be considered. This will depend largely upon the number of the observations and their distribution during the day, after the instrumental equipment has been determined upon.

164. Precision Desired and Number of Observations. In planning the observing program, having a definite end in view, it is advisable to decide upon the degree of precision which is to be sought in the result. This will depend to some extent upon the instruments or apparatus available, but, with a given instrument and an individual observer, the method of observing and the number of the observations become of great importance in deter-

[1] For a more extended treatment of this subject, the reader is referred to Wright and Hayford, Adjustment of Observations, Chapter IX.

[2] A notable instance of this was the design of the Coast and Geodetic Survey Precise Level, in 1900, by Mr. J. F. Hayford, Chief of the Computing Division, and Mr. E. G. Fischer, Chief of the Instrument Division, U. S. C. & G. S.

mining the precision. The observing program will frequently take the form of a number of units, or parts, all of which are alike with the exception of a change in the position of the instrument, as in the case of horizontal angles measured with a direction theodolite.

To attain the desired precision, then, the total number of observations must be considered. As a result of experience or experiment,[1] the precision (indicated by the mean square error, perhaps) of each elementary observation is ascertained, and from these, the precision of a unit observation. Then the number of observations necessary to obtain the desired precision in the result may be computed from the relation that the precision of the mean varies as the square root of the number of observations (Art. 141, page 163). That is, to double the precision (to divide the mean square or probable error by two) four times as many observations must be made. But how far can this process be continued? Is it possible to reach any degree of precision by simply multiplying the observations?

165. Ultimate Limit of Precision and Accuracy. While in theory the precision of the mean can be increased indefinitely by increasing the number of observations, experience shows that a limit is soon reached, beyond which it is not worth while to continue the observing; the theoretical increase in the precision as indicated by the smaller probable error, for example, would become quite misleading. Furthermore, after passing a certain point, the number of observations would have to be enormously increased in order to produce a very small decrease in the probable error, so that this process would be very wasteful of time and energy, and it is doubtful if the results would be much better.

After all, accuracy is desired rather than precision. The observations are not made for the purpose of enjoying the labor, but in order to ascertain the truth as far as practicable. It is a well-known fact that the mean of a small number of very consistent observations, showing a very small probable error, may be farther

[1] A theoretical discussion of the limitations of the human eye in making observations, and the increased power resulting from properly designed instruments, will be found in Jordan, Handbuch der Vermessungskunde, Band II, §45.

from the truth than that of a larger number of observations which vary over a considerable range. Cases can be cited in which a value adopted as a result of many observations, by different observers, extending over a long period of time, has been proved to be incorrect by an amount greater than many times the probable error. Of course, the conclusion is that we must not lose sight of the fact that, however consistent the observations may be, large systematic errors may be present and the observing methods may not be such as to eliminate them, so that they directly affect the results.

As to the limiting number of observations, then, we can safely state that this should be large enough and so distributed as to cover varying conditions as completely as practicable. Naturally, it will be different in various kinds of work. However, changes in the instrument and its supports are likely to take place if the observations extend over too much time, so that it is generally advisable to observe as rapidly as is practicable without a sacrifice of precision.

166. Indication of Systematic Errors. In order to discover the presence of systematic errors, a careful study of the residuals is essential. Unless the conditions causing these errors change during the course of the observations, the errors fall into the class of constant errors and will not be indicated at all by the discrepancies or residuals. In this case, a different method of observing might reveal them when the results of both methods were compared.

By plotting the residuals in chronological order some regularity or law may be noted in their occurrence. Positive and negative residuals may occur in separate groups or a curve drawn through the plotted points may show a periodic character. Again, the numbers of residuals of the various sizes may be plotted as in Art. 17, to form a Curve of Error, and if the resulting curve differs considerably at certain points from the theoretical form, which may be plotted from Table II or III in Appendix F, the presence of systematic errors may be indicated. Having thus investigated the occurrence of the residuals, it remains to seek changes in the observing conditions which correspond to the variations in the residuals. The location of such changes should serve to point out conditions responsible for part or all of the systematic errors so detected.

167. Treatment of Discordant Observations. When the discrepancies in a set of direct observations are unusually large, the lack of precision will be indicated by a large probable error or mean square error, and the mean remains as the best value obtainable from the given measures. It sometimes happens, however, that different *sets* of observations of the same quantity will yield results which are so discordant as to indicate the presence of constant or systematic errors in one or both of the sets. The problem may be further complicated by the fact that the precisions of the results may be considerably different, so that if their weighted mean were taken, as in Art. 156, it would give a decided preference to one of them. The question arises as to whether the results may not be so far apart as to make it advisable to neglect their relative weights altogether and to take their simple mean arbitrarily. This course is sometimes advocated.

Obviously, this is a matter of judgment rather than Least Squares, and such action should be preceded by a careful investigation of all the circumstances. However, it may be reasonably contended that if such discordant results are to be used at all a small difference in the adopted value would be of little moment and the regular Least Squares process may well be followed without considering the case as an exceptional one. Should conditions or checks be found which would be satisfied much more nearly by one of the results than by the other, the problem is thereby altered and becomes one involving the assignment of weights or perhaps the rejection of observations. The judgment of the computer must be the determining factor.

168. Arbitrary Adjustments. The principles outlined in the foregoing chapters, especially in Chapters V and VI, will be found of assistance in some problems where it may be deemed sufficient to approximate to a rigid adjustment by assigning corrections to the observed quantities arbitrarily. While such a method can hardly be defended in the hands of the computer who is conversant with Least Squares, still it must be admitted that such a computer is the only one who could be expected to carry out an arbitrary adjustment consistently and reasonably. The usual difficulty arises in satisfying all of the necessary conditions at the same time without a distribution of the corrections which is clearly unreasonable.

In certain problems, however, a distribution of arbitrary corrections may be of use in preparation for a rigid adjustment. The method consists in applying to the observations such preliminary corrections, resulting from a detailed study of the condition equations, as will reduce the amounts of the final corrections. This advance study requires a clear understanding of the field conditions as well as the methods of adjustment, but when carefully carried out is likely to diminish the labor of the computation and to improve the adjustment by reducing the numerical quantities involved. The method is analogous to the assumption of approximate values for the unknowns in the adjustment of indirect observations, Chapter III.

169. Use and Abuse of Least Squares. In view of the criticisms which are sometimes directed at the use of Least Squares for the adjustment of observations, a few words on the subject may not be out of place here. While it is unquestionably true that the method is sometimes used in an unwarranted manner, the real difficulty probably arises from the placing of erroneous interpretations upon, or the drawing of unreasonable conclusions from, the results of the adjustments.

A great deal of misunderstanding in the minds of persons unfamiliar with the fundamental principles of the method has resulted from the use of the term "probable error," and such persons are too apt to blame the method for the fruits of their misuse of it. It is unfortunate that this term has come into use, since its meaning in Least Squares is a technical one and not what would be expected from the ordinary use of the word "probable." Some of this trouble, to say the least, would have been avoided by using the "mean square error."

A common criticism relates to the use of Least Squares in connection with a very small number of observations, even as small as two. The reply may well be, "Where is a better method?"

The intelligent computer does not place the same reliance upon a very small number of observations as upon a larger one, but having only the small number he uses them as best he can. However, to place great confidence in the *precision* of the mean of two observations is certainly questionable, although even that precision may be very useful, in spite of its limitations, for purposes of comparison.

While investigations of the precision of observations and results have been thus criticised, little or no objection has ever been raised against the use of Least Squares for determining the best values of the unknown quantities. Its advantages for this purpose are evident even to those who are not familiar with its details. It provides a method of adjustment which is consistent, definite, and adaptable to the various kinds of problems and conditions, and which conforms to the facts as to the occurrence of errors of observation. Generally, also, it is simpler than an arbitrary adjustment; certainly it is more reliable.

170. Adjustments not Infallible. The beginner must not make the error of assuming that the results of an adjustment are correct. At the risk of repetition, this principle is emphasized, —that the results are but approximations to the true or correct values, the best obtainable from the given observations. Should the observations be affected by constant errors, the results will be likewise affected, without regard to their precision, which is determined from the discrepancies among them.

Also, as has been pointed out, different adjustments of the same observations by slightly different methods, perhaps, may yield results which are not exactly the same, owing to the fact that different sets of numerical quantities are used. If the computations are carried out to one decimal place more or less, slight variations in the final values may similarly occur. But it should be kept in mind that any one of these various adjustments will probably satisfy the requirements of the problem within the uncertainties among the observations, so that any one of them can safely be adopted.

171. Other Laws of Error. When applying the method of Least Squares to a new class of problems, it becomes necessary to investigate the occurrence of the errors, particularly when these are not actual errors of observation. It has been found by experiment that the variations among many natural occurrences follow the same law as the accidental errors of observation. Thus the law is applied in studies of the growth of vegetables, and to the occurrence of various characteristics among animals.

To illustrate errors which do not follow this law, we may consider the errors in a table of logarithms. It is evident that in a seven-place table, for example, the decimals following the seventh place have been rejected when less than 5 in the eighth place, while if the eighth place is greater than 5, the seventh place is increased by unity. Therefore, instead of the three assumptions upon which Least Squares is based (Art. 18), we have errors occurring only between the limits 0.0 and 0.5, the unit being in the last place of the logarithm, and in equal numbers without regard to magnitude or sign. The probabilities of the occurrence of the various errors between these limits would be equal, and the curve of error would be a rectangle upon the axis of errors as a base and limited by the ordinates at $+0.5$ and -0.5.

172. Review: Outline of Methods of Adjustment. In conclusion, a brief outline will be given covering the main classes of problems which have been considered and the methods of solution.

Direct Observations of a Single Quantity.

Adjustment. Take the mean or the weighted mean.

Indirect Observations.

Adjustment. Write the observation equations and from them the normal equations; the solution of the latter gives the unknown quantities themselves or the corrections to their assumed approximate values. The number of the observation equations will be the same as that of the observations; the number of the normal equations will equal that of the unknown quantities, which must always be less than that of the observations.

Conditioned Observations.

Adjustment. Write the condition equations in their general form and then in their simple form involving the corrections. From them form the normal equations, the same in number as the conditions. The solution of the normal equations gives a set of factors, called correlates, one for each condition equation, from which the desired corrections to the observed quantities are determined.

Simple Propagation of Error.

Solution. Write the literal function whose mean square error is desired. Differentiate it successively with respect to each of the quantities for which mean square errors are given. Substitute these partial derivatives and the given mean square errors in the general equation of propagation of error to obtain the mean square error of the function.

Compound Propagation of Error.

Solution. Find the mean square error of the function as above for each of the different sources of error, and take the square root of the sum of their squares.

Combination of Computed Quantities.

Adjustment. Give to each value a weight equal to the reciprocal of the square of its mean square error and take the weighted mean as the best value of the quantity.

Empirical Formulas.

Solution. Plot the observations and sketch a smooth curve through them. From this curve select the form of the desired equation. Write an observation equation of the selected form for each of the observations, reducing to the linear form if necessary. Write normal equations and solve them as in Indirect Observations, for the constants or coefficients of the formula.

APPENDIX A

HISTORY AND BIBLIOGRAPHY OF LEAST SQUARES

173. Historical Sketch.[1] The principle of the arithmetic mean is very old. But when the first indirect observations were made, probably in astronomy, the necessity for adjustment became apparent. Observation equations were written as early as 1748, by Euler. In 1757 Simpson stated the axiom that positive and negative errors occur with equal frequency, and in 1770 Lagrange considered the occurrence of errors from the standpoint of the theory of probability. Laplace, in 1774, in his "Mécanique Céleste," further investigated the subject and laid the foundation for the development of Least Squares.

It was not until the end of the 18th century, however, that the Method of Least Squares was introduced. The first publication of the principle of least squares was by Legendre, in 1805, in his "Nouvelles méthodes pour la détermination des orbites des comètes," and by him the name was given, "Méthode des moindres quarrés." Although there is no question as to the priority of publication, it seems well established that Gauss had actually developed and used the method itself since 1794, when he was a student at the University of Göttingen. His first publication on the subject, however, was not until 1809, in his classic work, "Theoria motus corporum cœlestium." But Gauss deserves more credit than anyone else for the further development of the Method of Least Squares, and as Merriman states,[2] "Few branches of science owe so large a proportion of subject-matter to the labors of one man."

The first publication of a theoretical derivation of the Law of Error was made by Dr. R. Adrian, of Reading, Pa., in 1808, in

[1] For more detailed information, the reader is referred to Jordan, Handbuch der Vermessungskunde, Band I, Einleitung.

[2] Merriman: Method of Least Squares.

the "Analyst or Mathematical Museum," at Philadelphia. Gauss published his in the next year, and various others have followed.

174. Growth of the Literature. The development of the subject is indicated by the rate at which publications devoted to it appeared. In 1877 Professor Merriman published an investigation [1] of the literature of Least Squares, as a result of which he deduced some interesting statistics. The following data are are based upon his work.

Prior to 1805, 22 titles were found. From that time on, averaging by decades, the rate of publication increased steadily from about two per year in 1810 to about ten per year in 1870. Altogether, 408 titles were listed up to 1875. Of these, 153 were published in Germany, 78 in France, 56 in Great Britain, and 34 in the United States, the remaining ones being scattered over eight countries. The German language was used in 167 instances, French in 110, and English in 90.

175. Bibliography. In addition to the paper by Merriman, referred to above, Gore's Bibliography of Geodesy, in the Report of the U. S. Coast and Geodetic Survey for 1887, will be found very useful in an investigation of the literature of this subject, although many important works have appeared since that time.

From the large number of books and parts of books devoted to Least Squares and the Adjustment of Observations, the following are selected for reference:

WRIGHT: Adjustment of Observations. Van Nostrand, New York, 1884. This is the classic work in the English language on this subject. The applications are principally geodetic. It has long been out of print, and was succeeded by

WRIGHT AND HAYFORD: Adjustment of Observations. Van Nostrand, 1907. Less comprehensive than the foregoing, but improved in many respects. Mainly geodetic.

JORDAN: Handbuch der Vermessungskunde, I. Metzler, Stuttgart, 1910. A very complete treatise, presented in a direct style which is easily read. Most valuable for reference. Geodetic.

HELMERT: Ausgleichungsrechnung. Teubner, Leipzig, 1907. Comprehensive and scholarly, but somewhat difficult to read. The notation is unusual. Geodetic and physical.

[1] Merriman: List of Writings Relating to the Method of Least Squares, published in the Transactions of the Connecticut Academy, New Haven, 1877.

KOLL: Methode der kleinsten Quadrate. Berlin, 1893. Extensive and practical with many applications.

CZUBER: Theorie der Beobachtungsfehler. Leipzig, 1891. Largely theoretical, with applications to life insurance and statistics.

MERRIMAN: Method of Least Squares. Wiley, New York, 1913. Geodetic applications but general in scope.

COMSTOCK: Method of Least Squares. Ginn, Boston, 1895. Astronomical and general.

BARTLETT: Method of Least Squares. Boston, 1915. Contains an extensive list of examples for solution.

WELD: Theory of Errors and Least Squares. Macmillan, New York, 1916. General and practical with many exercises for solution.

JOHNSON: Theory of Errors and Method of Least Squares. Wiley, 1892. General; strong in illuminating explanations.

BRUNT: Combination of Observations. Cambridge University Press, 1917. Theoretical.

CHAUVENET: Practical and Spherical Astronomy. Lippincott, Philadelphia, 1896.

CRANDALL: Geodesy and Least Squares. Wiley, New York, 1907.

ADAMS: Application of Least Squares to the Adjustment of Triangulation. Special Publication No. 28, U. S. C. & G. Survey, 1915. A very important contribution to this subject.

APPENDIX B

PRINCIPLES OF PROBABILITY

176. Definition. If an event can occur in a ways, and can fail to occur in b ways, the probability of its occurrence will be $\frac{a}{a+b}$, and that of its failure to occur will be $\frac{b}{a+b}$, it being assumed that all the ways of occurrence or failure to occur are *entirely independent* and *equally likely*. Thus, in one throw of a die, the probability of a certain face lying upward is $1/6$, and that of its not being upward, that is, of any other face being upward, is $5/6$. The probability of throwing *any* face upward will be $6/6=1$, in other words, certainty. Therefore, if the probability of the occurrence of an event be p, then that of the failure of the event to occur will be $1-p$, provided it is certain that the event must either occur or fail.

First Principle. The probability of the occurrence of an event is therefore a proper fraction between the limits zero (impossibility) and unity (certainty), and may be defined as the ratio between the number of ways in which the event may occur and the number of ways in which it may either occur or fail.

177. Two Sources of Probability. The probability of an event may be based upon theory or experience. The above case of throwing a die is an example of the theoretical basis. We know without question how many faces the die has and, therefore, the number of ways in which a certain face can lie upwards. The numbers involved are known absolutely. In the second case, on the other hand, the number of ways in which the event can occur is assumed as a result of experiment or experience. For example, if an event has occurred a times and failed b times out of a large number $a+b$, of trials, we may say that the probability of its occurrence, under the same conditions, is $a/(a+b)$, as before. Thus we may also define the probability of an event as the ratio of the number of times it has occurred to the total number of times it has occurred or failed; but the total number of cases, or attempts, should be sufficiently large to justify their use as a basis for generalization. To illustrate, suppose that statistics show that in the long run the number of male children born is to that of female children born as 21 to 20; then the probability that any birth will be that of a male is 21/41.

178. Simple Probability. The above statements, relating to the occurrence of a *single* event, illustrate simple probability. The principle will be further amplified. Suppose a box to contain w, white, b, black, and r red balls of the same weight and texture, and that a single ball is drawn from the box at random. Then the probability of drawing a ball of a certain color will be as follows:

White,	$\dfrac{w}{w+b+r}$
Black,	$\dfrac{b}{w+b+r}$
White or black,	$\dfrac{w+b}{w+b+r}$
Black or red,	$\dfrac{b+r}{w+b+r}$
White, black, or red,	$\dfrac{w+b+r}{w+b+r}=1$
Yellow,	$\dfrac{0}{w+b+r}=0$

Thus we may state the *Second Principle.* If the ways in which a single event can occur independently can be grouped in different sets or series, and the probability of its occurrence in each series be known, the total probability of its occurrence in *any* combination of the series will be the *sum* of the corresponding separate probabilities. In the above example, a single ball can be drawn from the white ones with a probability of

$$\frac{w}{w+b+r}$$

or from the black ones with a probability of

$$\frac{b}{w+b+r}$$

then the probability of drawing either a white or a black ball will be the sum of the two probabilities, namely, $\frac{w+b}{w+b+r}$.

179. Compound Probability. Independent Events. Suppose we have, in addition to the above box, a second one containing w' white, b' black, and r' red balls, and that we draw a ball from each box. Each of the $w+b+r$ possible draws from the first box may occur in combination with each of the $w'+b'+r'$ balls in the second so that the total number of possible draws of two balls, one from each box, will be $(w+b+r)(w'+b'+r')$. Also, each of the white balls in the first box may be drawn with each of the white ones in the second box, giving ww' possible pairs of white balls drawn one from each box. Therefore the probability of drawing simultaneously, two balls of one color, one from each box, will be,

$$\text{Two white balls, } \frac{ww'}{(w+b+r)(w'+b'+r')}$$

$$\text{Two red balls, } \frac{rr'}{(w+b+r)(w'+b'+r')}$$

$$\text{Two black balls, } \frac{bb'}{(w+b+r)(w'+b'+r')}$$

As a result of this reasoning, we can state the *Third Principle:* If two or more *independent* events are to occur simultaneously, and the probability of the separate occurrence of each is known, that of the simultaneous occurrence of all of them will be the *product* of the separate probabilities.

180. Compound Probability. Dependent Events. The probability of drawing a black and white pair, one ball from each of the two boxes, is an example in which the events are dependent. For, if a white ball were drawn from the first box, a black one would necessarily have to be drawn from the second box in order to make the pair, and vice versa, so that the probability of the second event would be different in the case of the failure of the first one than in its occurrence. Then the number of possible black and white pairs, one ball from each box, would be, $wb'+w'b$, and the probability of drawing such a pair would be,

$$\frac{wb'+w'b}{(w+b+r)(w'+b'+r')}$$

Here we have the occurrence of a compound event in two sets or series, so that the total probability is the sum of the separate (compound) probabilities.

Events are dependent when the probability of the occurrence of one of them depends upon the occurrence, or failure to occur, of another. By a careful analysis of each problem, it will usually be easy to so arrange or combine the events as to render them independent. In the foregoing example, the case of drawing a black and white pair, one ball from each box, is clearly one of dependent events, but if we require the probability of drawing a white ball from the first box simultaneously with a black one from the second box, the events are independent and, from the preceding article, the probability would be,

$$\frac{wb'}{(w+b+r)(w'+b'+r')}$$

Also, the probability of drawing a black ball from the first box and a white one from the second, simultaneously, would be

$$\frac{w'b}{(w+b+r)(w'+b'+r')}$$

But each of these events, while compound, is independent of the other. They are of the same character and may be considered as a single compound event occurring in two sets or series, so that the total probability of its occurrence in either manner will be, as in Art. 178, the sum of the two separate probabilities, that is,

$$\frac{wb'+w'b}{(w+b+r)(w'+b'+r')}$$

This is evident, also, from the first principle, when we note that the total number of black and white pairs is $wb'+w'b$, while the total number of possible pairs, of all colors, is

$$(w+b+r)(w'+b'+r')$$

181. Number of Occurrences. It follows from the definition of probability (Art. 176), that the number of times an event occurs may be determined by multiplying the total number of possible occurrences and failures, in other words, trials or attempts, by the probability of the occurrence of the event. In Least Squares, for example, the number of errors less than a certain amount to be expected in a given series of observations will be equal to the probability of an error less than that amount multiplied into the total number of observations in the set.

APPENDIX C

DERIVATION OF THE LAW OF ERROR

182. The Law of Error, that is, the equation of the Error Curve, (Art. 19), has been derived in several ways by different writers since the original demonstration by Dr. Adrian in 1808, published at Philadelphia in the "Analyst." The most notable of these, however, are the methods of Gauss (1809) and Hagen (1837). The former of these two will now be explained.[1]

183. Assumptions. The Error Function. Gauss based his derivation upon the assumption of the arithmetic mean as the most probable value of a directly observed quantity when all of the observations are made with the same care. Also, the occurrence of the errors of observation is assumed to be in accordance with the three axioms of Art. 18.

Since small errors are more numerous than large ones, and since the probability of an error of a certain size is directly proportional to the number of times that that error occurs in the given series of observations, it is evident that the probability of an error is a function of the error itself. Representing any error by Δ, the probability of the occurrence of this error by p_Δ, and the probability function by $\phi(\Delta)$, we can write,

$$p_\Delta = \phi(\Delta) \tag{308}$$

Strictly speaking, consecutive errors will differ by small finite amounts which are the least readings made with the given instrument or by the method used. For example, the least reading of a vernier on a circle may be $10''$, so that all the observations might be made only to the nearest $10''$, and the errors themselves

[1] See Brunt's Combination of Observations, for the methods of Hagen, Thomson and Tait, and Eddington. Hagen's proof is given in many works on Least Squares. Comstock frankly assumes the Law of Error to be empirical, which is a reasonable and practical method of attack.

would differ by multiples of 10″. So the ordinates to the error curve, corresponding to the various errors, and the successive points on the curve, would be separated by these intervals. However, as the precision of the observations increases, these elemental differences decrease and so we may reasonably regard the points as being so close together as to make the curve continuous. Thus we may consider that the errors, Δ, vary continuously, and that the function, $\phi(\Delta)$, is a continuous one. The probability of an error, Δ, is therefore equivalent to the probability of an error between the limits, Δ and $\Delta+d\Delta$.

The probability of the occurrence of an error between two limits is the sum of the separate probabilities of all the possible errors lying between those limits.[1] If we regard each probability as the ordinate to the error curve, corresponding to its particular error, the sum of these successive ordinates, when the curve is a continuous one, will constitute the area between the limiting ordinates, the curve, and the axis of Δ. Then, the probability of an error between Δ and $\Delta+d\Delta$ would be represented by the area of the infinitesimal vertical strip of length $\phi(\Delta)$ and of width $d\Delta$, that is, by the area $\phi(\Delta)d\Delta$. Therefore, the probability of the occurrence of an error between the limits a and b would be

$$\int_a^b \phi(\Delta)d\Delta$$

If the limits be extended so as to include all possible errors, namely, between $-\infty$ and $+\infty$, the probability of the occurrence of any error between these limits would be unity, that is, certainty, and this can be stated,

$$\int_{-\infty}^{+\infty} \phi(\Delta)d\Delta=1 \tag{309}$$

or, since the area is symmetrical about the axis of probability,

$$\int_0^{\infty} \phi(\Delta)d\Delta=\tfrac{1}{2} \tag{310}$$

[1] See Appendix B, Principles of Probability.

184. Derivation of the Law of Error. We shall consider the general case of indirect observations, since direct observations form but a special case under it. The observed quantity is a function of the unknown quantities. Let there be n observations and m unknowns, n being greater than m (Art. 22). The observation equations may be written,

$$\begin{aligned} f_1(X, Y, Z, \quad . \quad . \quad .) &= M_1 \\ f_2(X, Y, Z, \quad . \quad . \quad .) &= M_2 \\ . \quad . \quad . \quad . \quad . \quad . \quad &. \quad . \quad . \quad . \quad . \\ f_n(X, Y, Z, \quad . \quad . \quad .) &= M_n \end{aligned} \tag{311}$$

Let $\Delta_1, \Delta_2, \Delta_3, \ . \ . \ . \ \Delta_n$, be the respective errors of $M_1, M_2, M_3, \ . \ . \ . \ M_n$, and let the probability of the occurrence of Δ_1 be $\phi(\Delta_1)$, that of Δ_2 be $\phi(\Delta_2)$, etc. Then the probability of the simultaneous occurrence of this series of errors will be the product of their separate probabilities, or,

$$P = \phi(\Delta_1)\phi(\Delta_2)\phi(\Delta_3) \quad . \quad . \quad . \quad \phi(\Delta_n) \tag{312}$$

Taking the logarithm of each member, this becomes,

$$\log P = \log \phi(\Delta_1) + \log \phi(\Delta_2) + \quad . \quad . \quad . \quad + \log \phi(\Delta_n) \tag{313}$$

The most probable series of errors will be those for which the above probability is a maximum, which also will be the case when $\log P$ is a maximum. This is the condition for the best or most probable values of the unknowns. Therefore, the first derivative of (313) must equal zero, and since the unknowns $X, Y, Z, \ . \ . \ .$, in the case of indirect observations are independent, it follows that the separate partial derivatives of $\log P$ with respect to these unknowns must equal zero. Thus we obtain,

$$\begin{aligned} \frac{d\phi(\Delta_1)}{\phi(\Delta_1)dX} + \frac{d\phi(\Delta_2)}{\phi(\Delta_2)dX} + \quad . \quad . \quad . \quad + \frac{d\phi(\Delta_n)}{\phi(\Delta_n)dX} &= 0 \\ \frac{d\phi(\Delta_1)}{\phi(\Delta_1)dY} + \frac{d\phi(\Delta_2)}{\phi(\Delta_2)dY} + \quad . \quad . \quad . \quad + \frac{d\phi(\Delta_n)}{\phi(\Delta_n)dY} &= 0 \\ . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad &. \quad . \quad . \end{aligned} \tag{314}$$

Multiplying and dividing each fraction by the corresponding $d\Delta$,

$$\frac{d\phi(\Delta_1)}{\phi(\Delta_1)d\Delta_1}\ \frac{d\Delta_1}{dX}+\frac{d\phi(\Delta_2)}{\phi(\Delta_2)d\Delta_2}\ \frac{d\Delta_2}{dX}+\ \ldots\ +\frac{d\phi(\Delta_n)}{\phi(\Delta_n)d\Delta_n}\ \frac{d\Delta_n}{dX}=0$$

$$\frac{d\phi(\Delta_1)}{\phi(\Delta_1)d\Delta_1}\ \frac{d\Delta_1}{dY}+\frac{d\phi(\Delta_2)}{\phi(\Delta_2)d\Delta_2}\ \frac{d\Delta_2}{dY}+\ \ldots\ +\frac{d\phi(\Delta_n)}{\phi(\Delta_n)d\Delta_n}\ \frac{d\Delta_n}{dY}=0 \qquad (315)$$

$$\ldots\ldots\ldots\ldots\ldots\ldots$$

Since the function, $\phi(\Delta)d\Delta$, must be applicable to any number of unknowns, we shall make use of the case of one unknown, directly observed, from which to determine the nature of the function. Letting X represent the true value of the unknown, and Δ the true error of the observation, M, we may write,

$$X-M_1=\Delta_1$$

$$X-M_2=\Delta_2 \qquad (316)$$

$$\ldots\ldots$$

$$X-M_n=\Delta_n$$

Differentiating,

$$\frac{d\Delta_1}{dX}=\frac{d\Delta_2}{dX}=\frac{d\Delta_3}{dX}=\ \ldots\ =\frac{d\Delta_n}{dX}=1 \qquad (317)$$

Substituting in (315),

$$\frac{d\phi(\Delta_1)}{\phi(\Delta_1)d\Delta_1}+\frac{d\phi(\Delta_2)}{\phi(\Delta_2)d\Delta_2}+\ \ldots\ +\frac{d\phi(\Delta_n)}{\phi(\Delta_n)d\Delta_n}=0 \qquad (318)$$

Multiplying and dividing each fraction by the corresponding Δ,

$$\frac{d\phi(\Delta_1)}{\Delta_1\phi(\Delta_1)d\Delta_1}\Delta_1+\frac{d\phi(\Delta_2)}{\Delta_2\phi(\Delta_2)d\Delta_2}\Delta_2+\ \ldots\ +\frac{d\phi(\Delta_n)}{\Delta_n\phi(\Delta_n)d\Delta_n}\Delta_n=0 \qquad (319)$$

But it is assumed in direct observations that the mean is the best or most probable value of the observed quantity, and that, as the number of observations increases indefinitely, the mean approaches the true value as a limit. So we may write,

$$X=\frac{M_1+M_2+\ \ldots\ +M_n}{n} \qquad (320)$$

or,

$$(X-M_1)+(X-M_2)+\ \ldots\ +(X-M_n)=0 \qquad (321)$$

whence, from (316),

$$\Delta_1+\Delta_2+\ .\ .\ .\ +\Delta_n=0 \tag{322}$$

Both (319) and (322) hold good as the number of observations is increased one by one. But in order that this condition may exist, it is necessary that the coefficients of the Δ's in (319) be equal and constant, so that they may be cancelled from that equation.

Therefore, we may write, in general,

$$\frac{d\phi(\Delta)}{\Delta\phi(\Delta)d\Delta}=\text{a constant, say } k \tag{323}$$

or,

$$\frac{d\phi(\Delta)}{\phi(\Delta)d\Delta}=k\Delta \tag{324}$$

Integrating,

$$\log\ \phi(\Delta)=\tfrac{1}{2}k\Delta^2+k' \tag{325}$$

whence,[1]

$$\phi(\Delta)=e^{\frac{1}{2}k\Delta^2}e^{k'} \tag{326}$$

But one of the original assumptions was that small errors are more numerous and more probable than large ones. Thus, as Δ decreases, $\phi(\Delta)$ must increase, which requires that k must always be negative. To effect this, we replace $k/2$ by the new constant, $-h^2$. Then, replacing the constant factor, $e^{k'}$, by the constant, C, we obtain the expression for the probability of an error, Δ,

$$\phi(\Delta)=Ce^{-h^2\Delta^2} \tag{327}$$

185. The Constant, C. It remains to determine the value of the constant, C. Substituting (327) in (310),

$$\int_0^\infty Ce^{-h^2\Delta^2}d\Delta=\tfrac{1}{2} \tag{328}$$

Let $t=h\Delta$; then $dt=hd\Delta$. Also, when $\Delta=0$, $t=0$, and when $\Delta=\infty$, $t=\infty$. Therefore we may write (328) as follows,

$$\frac{C}{h}\int_0^\infty e^{-h^2\Delta^2}hd\Delta=\tfrac{1}{2} \tag{329}$$

[1] e is the base of Napierian logarithms.

$$\int_0^\infty e^{-t^2}d = \frac{h}{2C} \tag{330}$$

But,[1]

$$\int^\infty e^{-t^2}dt = \frac{\sqrt{\pi}}{2} \tag{331}$$

so that, from (330) and (331),

$$\frac{h}{2C} = \frac{\sqrt{\pi}}{2}$$

whence,

$$C = \frac{h}{\sqrt{\pi}} \tag{332}$$

Therefore, we have from (327) the final expression for the Law of Error,

$$\phi(\Delta) = \frac{h}{\sqrt{\pi}}e^{-h^2\Delta^2} \tag{333}$$

[1] This definite integral may be evaluated in various ways. The following method is given by Bartlett:

From the assumption, $t = h\Delta$, we have,

$$\int_0^\infty e^{-t^2}dt = \int_0^\infty e^{-h^2\Delta^2}h\,d\Delta \tag{334}$$

But when definite integrals, only, are used,

$$\int_0^\infty e^{-t^2}dt = \int_0^\infty e^{-h^2}dh \tag{335}$$

Multiplying (334) and (335),

$$\left[\int_0^\infty e^{-t^2}dt\right]^2 = \int_0^\infty \int_0^\infty e^{-h^2(1+\Delta^2)}h\,d\Delta\,dh \tag{336}$$

$$= \int_0^\infty \frac{-d\Delta}{2(1+\Delta^2)} \int_0^\infty e^{-h^2(1+\Delta^2)}(-2h)(1+\Delta^2)dh \tag{337}$$

$$= \int_0^\infty \frac{-d\Delta}{2(1+\Delta^2)} \left[e^{-h^2(1+\Delta^2)}\right]_0^\infty \tag{338}$$

$$= \frac{1}{2}\int_0^\infty \frac{d\Delta}{1+\Delta^2} \tag{339}$$

$$= \frac{1}{2}\left[\tan^{-1}\Delta\right]_0^\infty = \frac{\pi}{4} \tag{340}$$

Therefore,

$$\int_0^\infty e^{-t^2}dt = \frac{\sqrt{\pi}}{2} \tag{341}$$

186. Expansion of Law of Error in Series. The Law of Error may be expressed in the form of series for convenience in evaluating it for various values of Δ.[1]

Using the quantity, $t=h\Delta$, as an auxiliary variable, we can state (333) as follows:

$$\phi(\Delta)d\Delta=\frac{1}{\sqrt{\pi}}e^{-h^2\Delta^2}hd\Delta=\frac{1}{\sqrt{\pi}}e^{-t^2}dt \tag{342}$$

which is the probability of an error between Δ and $\Delta+d\Delta$. The probability of the occurrence of an error less than Δ will be that of an error between the limits, $-\Delta$ and $+\Delta$, that is, since $t=h\Delta$,

$$p_{-\Delta}^{+\Delta}=\int_{-\Delta}^{+\Delta}\phi(\Delta)d\Delta=\frac{1}{\sqrt{\pi}}\int_{-t}^{+t}e^{-t^2}dt=\frac{2}{\sqrt{\pi}}\int^{+t}e^{-t^2}dt \tag{343}$$

$$=\frac{2}{\sqrt{\pi}}\left(t-\frac{t^3}{3\cdot 1!}+\frac{t^5}{5\cdot 2!}-\frac{t^7}{7\cdot 3!}\quad\cdot\ \cdot\ \cdot\right) \tag{344}$$

for use with small values of t,

or,

$$=1-\frac{e^{t^2}}{t\sqrt{\pi}}\left(1-\frac{1}{2t^2}+\frac{1\cdot 3}{(2t^2)^2}-\frac{1\cdot 3\cdot 5}{(2t^2)^3}+\quad\cdot\ \cdot\ \cdot\right) \tag{345}$$

for use with large values of t.

187. Tables of the Law of Error. From the above formulas, tables have been computed with the argument t, giving the probability of an error less than Δ, in a given set of observations. Table I, in Appendix F, has been formed in this manner. To use such a table, the mean square error of a single observation, (ϵ), is computed from the residuals of the mean. Then t is obtained from the assumed error, Δ, by means of the relation,

$$t=h\Delta=\frac{\Delta}{\epsilon\sqrt{2}} \tag{346}$$

since, from (209),

$$h=\frac{1}{\epsilon\sqrt{2}}$$

Finally, with t as an argument, the tabular probability is obtained.

[1] See Wright and Hayford, Art. 22; Crandall, Art. 141; or Chauvenet, Vol. I, Art. 113.

However, it is more convenient in many cases to express the function in terms of Δ/ϵ or Δ/r directly, and this has been done in Tables II and III, respectively, in Appendix F. The table gives the probability of an error less than a certain fraction (Δ/ϵ) of the mean square error, or (Δ/r), of the probable error, of a single observation. Thus, from Table II, the probability of an error less than 0.4 of the mean square error is 0.3108, and the number of such errors should be approximately 0.3108 times the number of observations, n, in the given series. Similarly, the number of errors greater than 0.4ϵ would theoretically be, $n(1-0.3108)$. By comparing these theoretical numbers of errors with those actually counted in the given set, it is possible to ascertain how closely the observations conform to the theory (Art. 20).

APPENDIX D

OUTLINE OF A SHORT COURSE OF INSTRUCTION

188. General Plan. While it is desirable to devote a three-hour course for one semester to the study and practice of Least Squares and the Adjustment of Observations, with civil engineering students, the author presents the following outline of a *one-hour course* which he conducted at Cornell University when, owing to the demands of other courses, this was all the time which the student could devote to the subject. He regards such a course as very much worth while and believes that the students obtained a good general knowledge of the methods of adjusting observations together with considerable practice in the solution of problems.

The course was given in 16 lessons, and in addition to the fifty-minute lecture, the student was expected to work two hours at home upon the text and the assigned problem. The problems were handed in at the next lecture, with a penalty for failing to do so. It was considered essential that the problem be solved while the topic was fresh in the student's mind. The problems were carefully examined by comparison with standards and returned for correction, if necessary, or retained until the end of the term. The work was required to be neatly done with the idea that the set of examples would be kept for reference.[1]

The lectures had to be limited to the essential parts of the subject, owing to the limited time, and especial attention was given to the solution of the problem at hand. Sometimes two lectures intervened between problems, and in the case of the double problem of the adjustment of a quadrilateral a lecture was omitted in order to give the student more time for the solution. The

[1] The paper known as "Data Sheets" was used. It is $8 \times 10\frac{1}{2}$ inches and ruled with blue lines one-fourth inch apart parallel to the shorter edge and with perpendicular red lines forming ten equal columns. A blank margin is left at the top and left-hand edges of the sheet.

first lectures were devoted to a very careful consideration of the occurrence of errors. Thence the order is indicated by the problems in the following list.

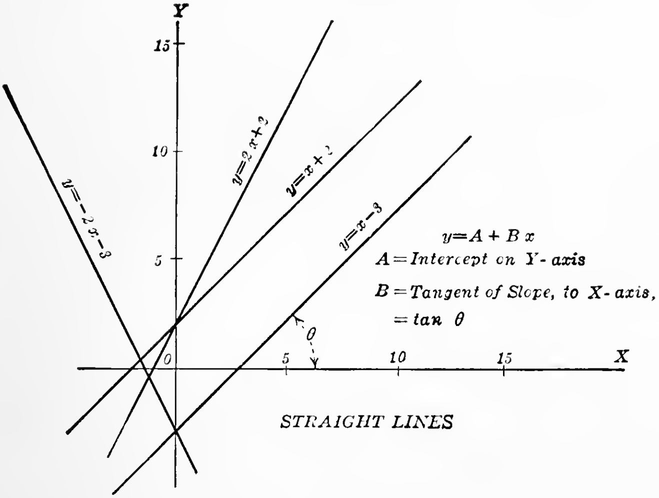
189. List of Problems. It was intended that each of the ordinary problems would be of such length that the average student could solve it in two hours, in connection with the accompanying text. The order here given may be varied, if desired, and Nos. 9 and 10 may be combined. The inclusion of the topic of index of precision and mean square error in the introductory lectures will depend upon the preference of the instructor; it is not necessary to introduce it until the propagation of error is to be studied.

1. Simple and weighted means; precision and mean square error.
2. Indirect observations; observation equations given; direct solution for the unknowns.
3. Indirect observations; observation equations given; solution with approximate values of unknowns to find corrections to those values.
4. Local adjustment of angles at a station.
5. Local adjustment; method of directions.
6. Adjustment of a level net.
7. Adjustment of a quadrilateral; method of angles.
8. Adjustment of a quadrilateral; method of directions. (Problems 7 and 8 may be combined, using same data.)
9. Simple propagation of error.
10. Compound propagation of error.
11. Combination of results; weights from mean square errors.

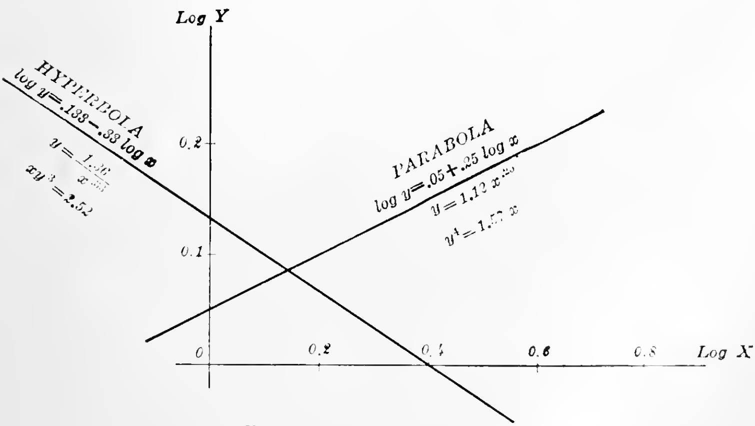
APPENDIX E

TYPICAL CURVES FOR REFERENCE

PLATE I



STRAIGHT LINES



NOTE: See next pages for these curves plotted by x and y

PLATE II

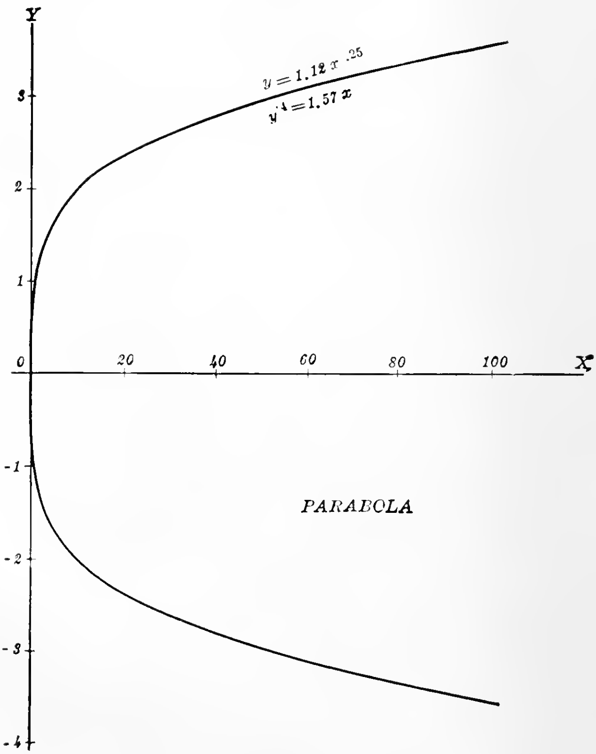
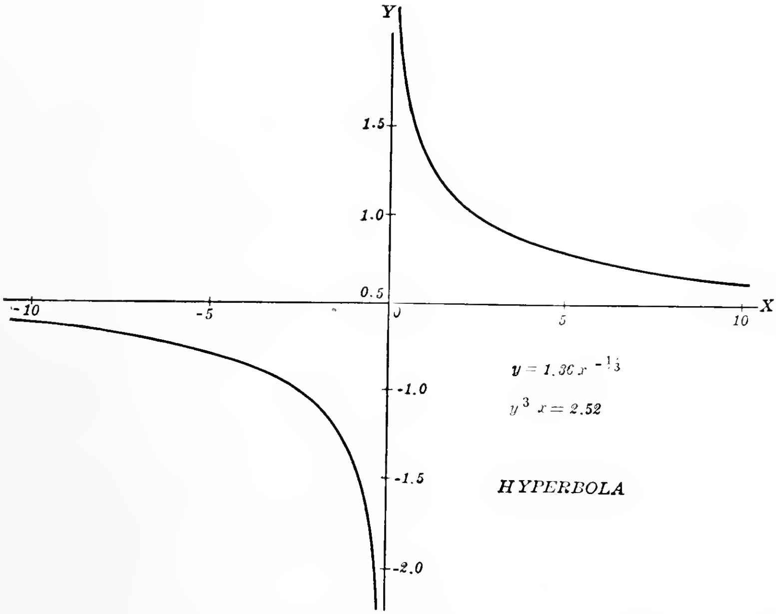


PLATE III


Y
X
1.5
1.0
0.5
-10
-5
0
5
10
-1.0
-1.5
-2.0
$y = 1.36\,x^{-1/3}$
$y^3 x = 2.52$
HYPERBOLA

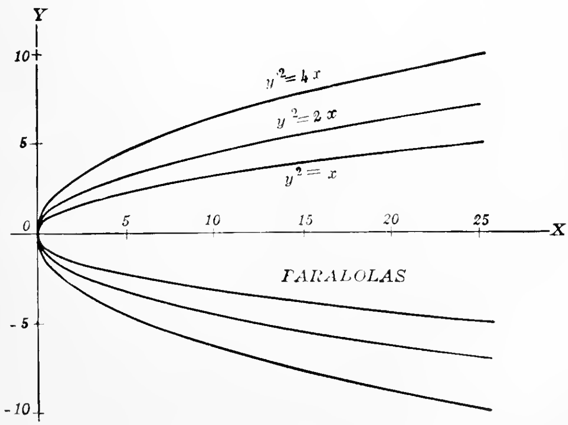

Y
X
10
5
0
-5
-10
5
10
15
20
25
$y^2 = 4x$
$y^2 = 2x$
$y^2 = x$
PARABOLAS

PLATE IV

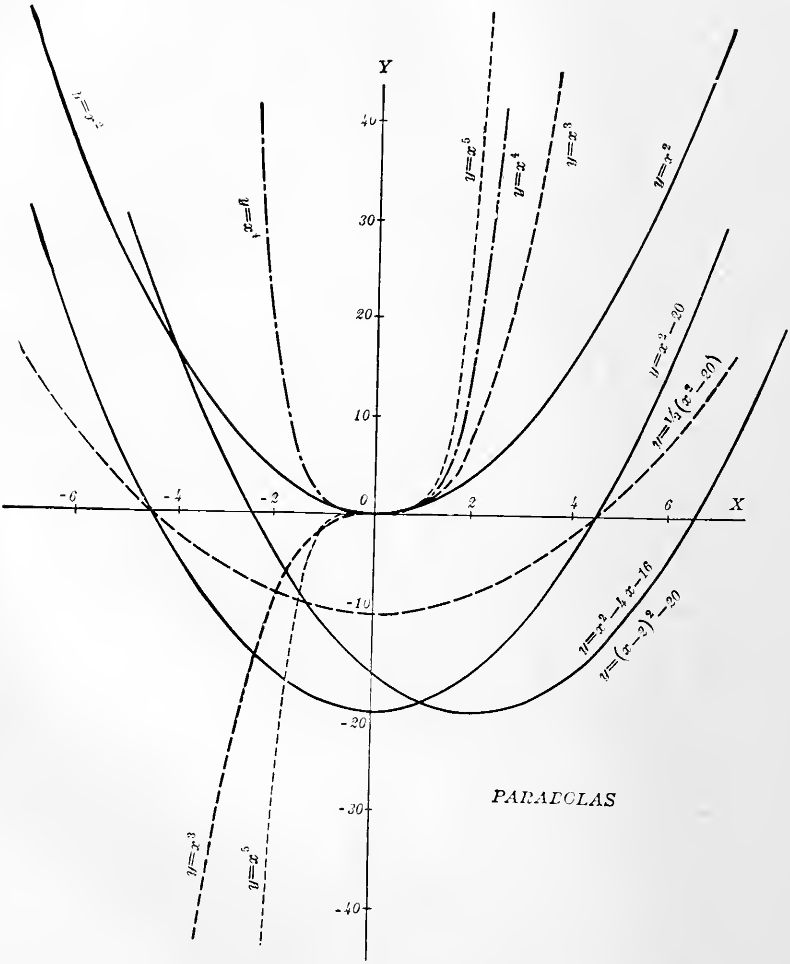

Y
X
40
30
20
10
-10
-20
-30
-40
-6
-4
-2
0
2
4
6
$y=x^2$
$x=y^4$
$y=x^5$
$y=x^4$
$y=x^3$
$y=x^2$
$y=x^2-20$
$y=\frac{1}{2}(x^2-20)$
$y=x^2-4x-16$
$y=(x-2)^2-20$
PARABOLAS
$y=x^3$
$y=x^5$

PLATE V

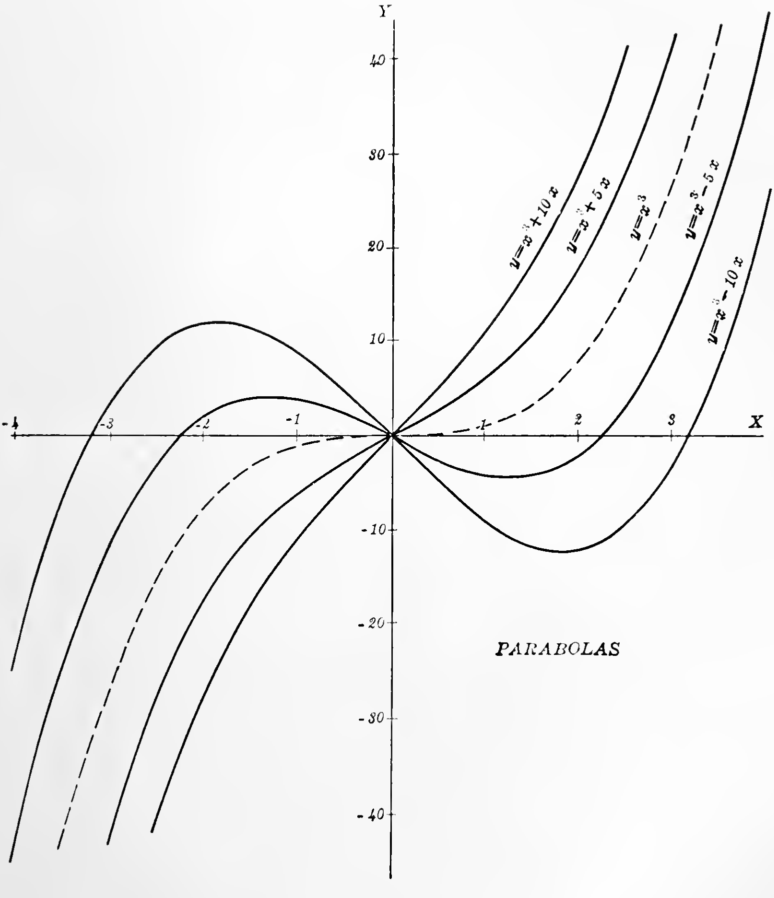

Y
X
40
30
20
10
-10
-20
-30
-40
-4
-3
-2
-1
1
2
3
$y=x^3+10x$
$y=x^3+5x$
$y=x^3$
$y=x^3-5x$
$y=x^3-10x$
PARABOLAS

PLATE VI

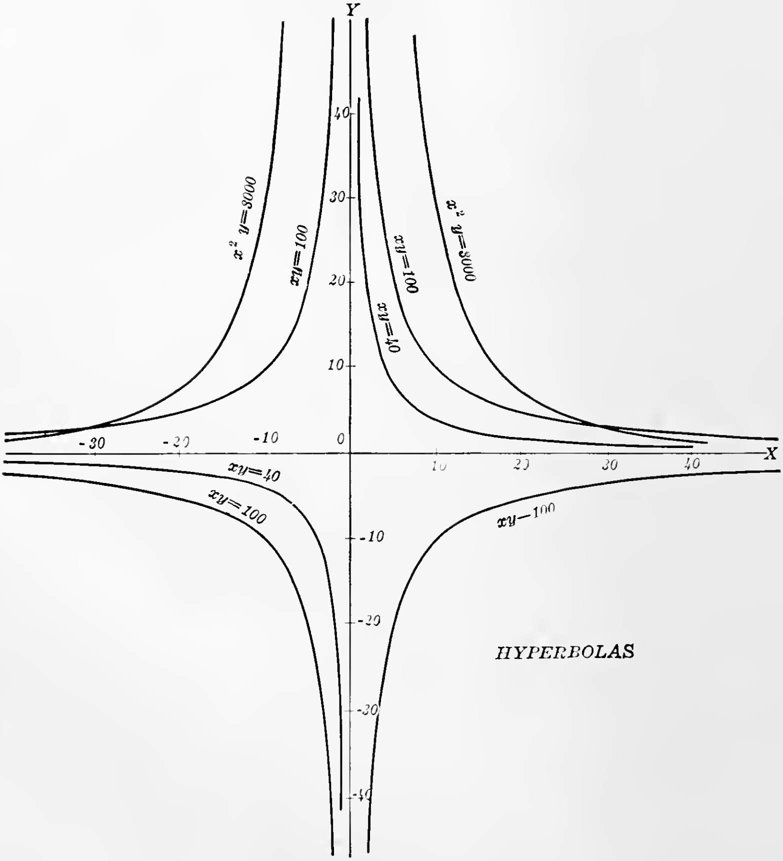

Y
X
40
30
20
10
0
-10
-20
-30
-40
- 30
- 20
10
20
30
40
$x^2 y=3000$
$xy=100$
$xy=40$
$xy=100$
$x^2 y=3000$
$xy=40$
$xy=100$
$xy=100$
HYPERBOLAS

PLATE VII

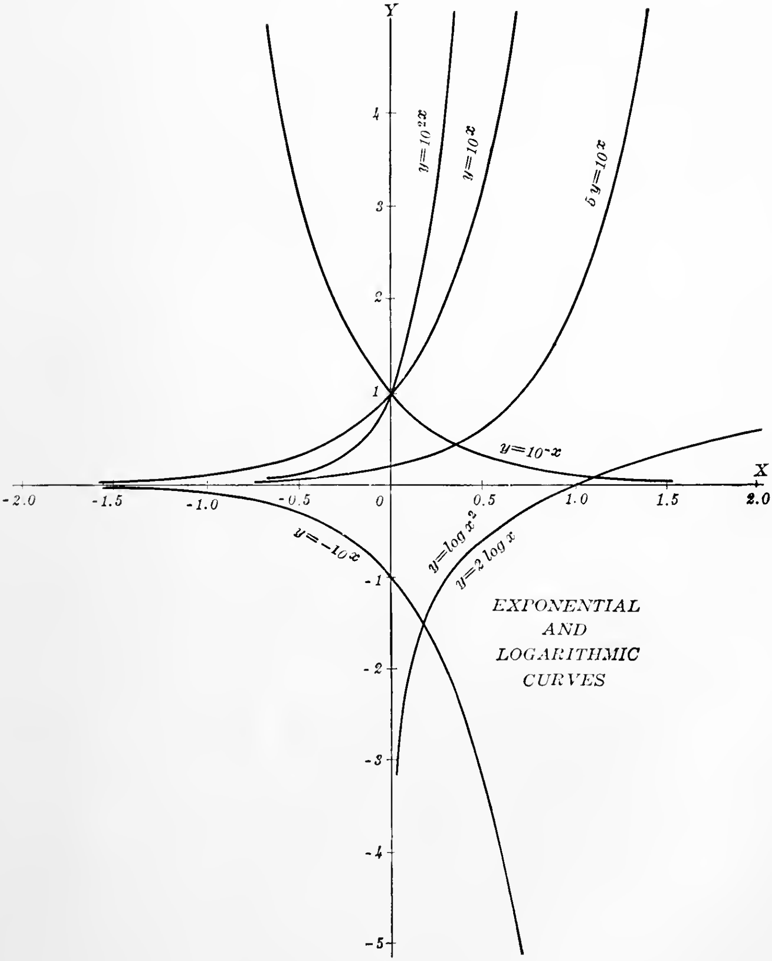
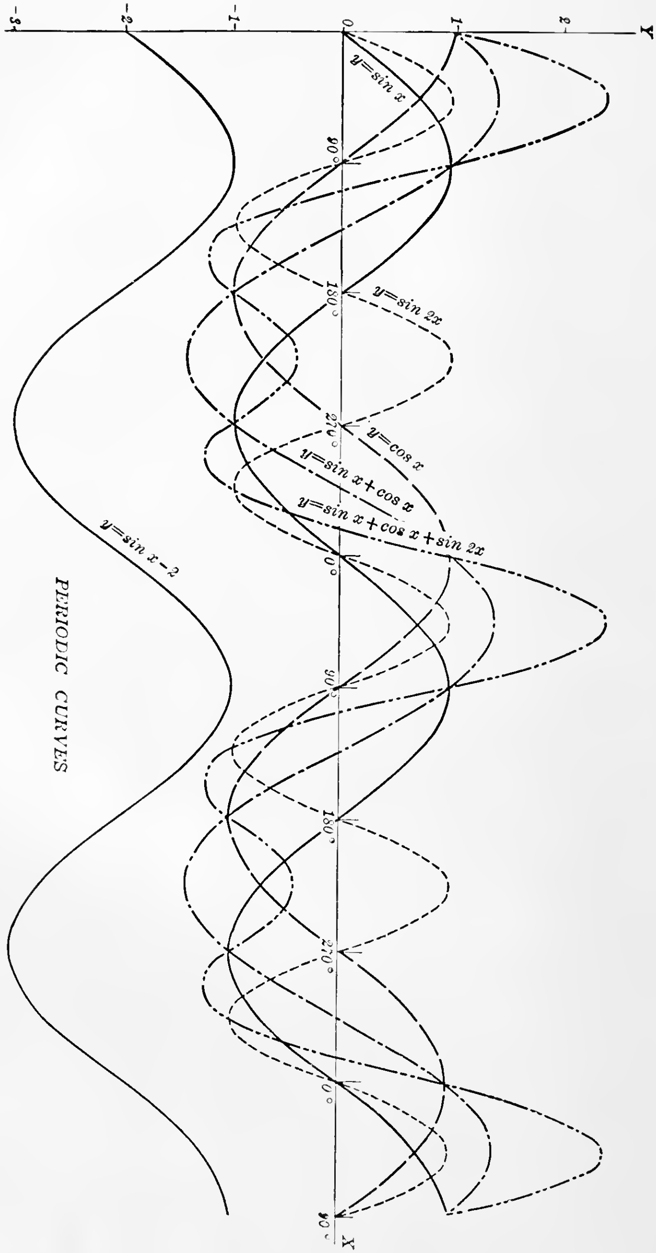

Y
X
y=10^2x
y=10^x
5y=10^x
y=10^-x
y=-10^x
y=log x^2
y=2 log x
4
3
2
1
0
-1
-2
-3
-4
-5
-2.0
-1.5
-1.0
-0.5
0.5
1.0
1.5
2.0
EXPONENTIAL
AND
LOGARITHMIC
CURVES

PLATE VIII



PERIODIC CURVES

APPENDIX F

TABLES

TABLE I

Values of $p = \frac{2}{\sqrt{\pi}} \int_0^t e^{-t^2} dt$

(Arts. 136 and 187) Argument is $t = h\Delta$

t	p	t	p
0.00	0.0000	1.00	0.8427
0.05	.0564	1.05	.8624
0.10	.1125	1.10	.8802
0.15	.1680	1.15	.8961
0.20	.2227	1.20	.9103
0.25	0.2763	1.25	0.9229
0.30	.3286	1.30	.9340
0.35	.3794	1.35	.9438
0.40	.4284	1.40	.9523
0.45	.4755	1.45	.9597
0.50	0.5205	1.50	0.9661
0.55	.5633	1.60	.9763
0.60	.6039	1.70	.9838
0.65	.6420	1.80	.9891
0.70	.6778	1.90	.9928
0.75	0.7112	2.00	0.9953
0.80	.7421	2.10	.9970
0.85	.7707	2.20	.9981
0.90	.7969	2.30	.9989
0.95	.8209	2.40	.9993
1.00	0.8427	2.50	0.9996

TABLE II

Values of $p=\dfrac{2h}{\sqrt{\pi}}\displaystyle\int_0^{\Delta} e^{-h^2\Delta^2}d\Delta$ in terms of $\dfrac{\Delta}{\epsilon}$

Probability of the occurrence of an error less than Δ.

$\epsilon=\sqrt{\dfrac{[v^2]}{n-1}}$, that is, the mean square error of a single observation.

(Art. 187, p. 215)

$\frac{\Delta}{\epsilon}$	p	$\frac{\Delta}{\epsilon}$	p
0.0	0.0000	2.0	0.9545
0.1	.0797	2.1	.9643
0.2	.1585	2.2	.9722
0.3	.2358	2.3	.9785
0.4	.3108	2.4	.9836
0.5	0.3829	2.5	0.9876
0.6	.4515	2.6	.9907
0.7	.5161	2.7	.9931
0.8	.5763	2.8	.9949
0.9	.6319	2.9	.9963
1.0	0.6827	3.0	0.9973
1.1	.7287	3.1	.9981
1.2	.7699	3.2	.9986
1.3	.8064	3.3	.9990
1.4	.8385	3.4	.9993
1.5	0.8664	3.5	0.9995
1.6	.8904	3.6	.9997
1.7	.9109	3.7	.9998
1.8	.9281	3.8	.9999
1.9	.9426	3.9	.9999
2.0	0.9545	4.0	0.9999
		4.1	1.0000

TABLE III

Values of $p = \frac{2h}{\sqrt{\pi}}\int_0^{\Delta} e^{-h^2\Delta^2}d\Delta$ in terms of $\frac{\Delta}{r}$

Probability of the occurrence of an error less than Δ.

$r = 0.6745\sqrt{\frac{[v^2]}{n-1}}$ = the probable error of a single observation.

(Art. 187, p. 215)

$\frac{\Delta}{r}$	p	$\frac{\Delta}{r}$	p
0.0	0.0000	2.5	0.9082
0.1	.0538	2.6	.9205
0.2	.1073	2.7	.9314
0.3	.1603	2.8	.9410
0.4	.2127	2.9	.9495
0.5	0.2641	3.0	0.9570
0.6	.3143	3.1	.9635
0.7	.3632	3.2	.9691
0.8	.4105	3.3	.9740
0.9	.4562	3.4	.9782
1.0	0.5000	3.5	0.9818
1.1	.5419	3.6	.9848
1.2	.5817	3.7	.9874
1.3	.6194	3.8	.9896
1.4	.6550	3.9	.9915
1.5	0.6883	4.0	0.9930
1.6	.7195	4.1	.9943
1.7	.7485	4.2	.9954
1.8	.7753	4.3	.9963
1.9	.8000	4.4	.9970
2.0	0.8227	4.5	0.9976
2.1	.8433	4.6	.9981
2.2	.8622	4.7	.9985
2.3	.8792	4.8	.9988
2.4	.8945	4.9	.9991
2.5	0.9082	5.0	0.9993

TABLE IV

Factors for Computing Probable Errors from Bessel's Formulas.

(Arts. 140 and 141)

n	$\frac{0.6745}{\sqrt{n-1}}$	$\frac{0.6745}{\sqrt{n(n-1)}}$	n	$\frac{0.6745}{\sqrt{n-1}}$	$\frac{0.6745}{\sqrt{n(n-1)}}$
			30	0.1252	0.0229
			31	.1231	.0221
2	0.6745	0.4769	32	.1211	.0214
3	.4769	.2754	33	.1192	.0208
4	.3894	.1947	34	.1174	.0201
5	0.3372	0.1508	35	0.1157	0.0196
6	.3016	.1231	36	.1140	.0190
7	.2754	.1041	37	.1124	.0185
8	.2549	.0901	38	.1109	.0180
9	.2385	.0795	39	.1094	.0175
10	0.2248	0.0711	40	0.1080	0.0171
11	.2133	.0643	41	.1066	.0167
12	.2029	.0587	42	.1053	.0163
13	.1947	.0540	43	.1041	.0159
14	.1871	.0500	44	.1029	.0155
15	0.1803	0.0465	45	0.1017	0.0152
16	.1742	.0435	46	.1005	.0148
17	.1686	.0409	47	.0994	.0145
18	.1636	.0386	48	.0984	.0142
19	.1590	.0365	49	.0974	.0139
20	0.1547	0.0346	50	0.0964	0.0136
21	.1508	.0329	51	.0954	.0134
22	.1472	.0314	52	.0944	.0131
23	.1438	.0300	53	.0935	.0128
24	.1406	.0287	54	.0926	.0126
25	0.1377	0.0275	55	0.0918	0.0124
26	.1349	.0265	56	.0909	.0122
27	.1323	.0255	57	.0901	.0119
28	.1298	.0245	58	.0893	.0117
29	.1275	.0237	59	.0886	.0115
30	0.1252	0.0229	60	0.0878	0.0113

INDEX

(Numbers refer to pages)

3.00