PUBLICATIONS OF THE NATIONAL BUREAU OF
ECONOMIC RESEARCH, INCORPORATED

No. 19

THE SMOOTHING OF TIME SERIES

National Bureau of Economic Research, Inc.

Publications in the order of their issue

INCOME IN THE UNITED STATES
By Wesley C. Mitchell, Willford I. King, Frederick R. Macaulay and Oswald W. Knauth
Volume I. A summary of an Investigation of the Amount and Distribution of Income in the United States, 1909-1919. 152 pages. $1.58.
Volume II. Details. 440 pages. $5.15.

DISTRIBUTION OF INCOME BY STATES IN 1919
By Oswald W. Knauth
The share of each state in the national income. 32 pages. (Out of print.)

BUSINESS CYCLES AND UNEMPLOYMENT
By the staff of the Bureau with 16 collaborators.
405 pages. $4.10.

EMPLOYMENT, HOURS, AND EARNINGS IN PROSPERITY AND DEPRESSION
By Willford I. King
147 pages. (Out of print.)

THE GROWTH OF AMERICAN TRADE UNIONS, 1880-1923
By Leo Wolman
Membership year by year; fluctuations with business cycles. 170 pages, $3.

INCOME IN THE VARIOUS STATES
By Maurice Leven and Willford I. King
Gives income for each state, in 1919, 1920 and 1921. 306 pages, $5.

BUSINESS ANNALS
By Willard L. Thorp and Wesley C. Mitchell
Covers 17 countries, 36 to 136 years. 384 pages, $4.

MIGRATION AND BUSINESS CYCLES
By Harry Jerome
A century of American experience. 256 pages, $3.50.

BUSINESS CYCLES: THE PROBLEM AND ITS SETTING
By Wesley C. Mitchell
Here Dr. Mitchell explains the numerous processes involved in business cycles; shows how our present economic organization was evolved; describes how to use statistics in the study of business cycles and how to use business annals; and concludes with a working concept of business cycles. (Eighth printing.) 512 pages, $6.50.

THE BEHAVIOR OF PRICES
By Frederick C. Mills
A pioneer study of individual commodity prices. 598 pages, $7.

TRENDS IN PHILANTHROPY
By Willford I. King
A study of a typical American city. 78 pages, $1.75.

RECENT ECONOMIC CHANGES
By Edwin F. Gay, Wesley C. Mitchell, and others
Conditions in the United States, 1922-1928. 2 volumes, $7.50 per set.

INTERNATIONAL MIGRATIONS
Vol. I. Compiled by Imre Ferenczi, International Labour Office. 1,112 pages, $10.
Vol. II. Edited by Walter F. Willcox. 700 pages, $7.

THE NATIONAL INCOME AND ITS PURCHASING POWER
By Willford I. King. 394 pages, $5.

CORPORATION CONTRIBUTIONS TO ORGANIZED COMMUNITY WELFARE SERVICES
By Pierce Williams and Frederick E. Croxton 347 pages, $3.

PLANNING AND CONTROL OF PUBLIC WORKS
Report of the Committee on Recent Economic Changes, and the report of Leo Wolman. 294 pages, $3.

THE SMOOTHING OF TIME SERIES
By Frederick R. Macaulay $2.

Copies of available reports may be obtained upon application accompanied by remittance to

NATIONAL BUREAU OF ECONOMIC RESEARCH, Inc.
51 Madison Avenue, New York

National Bureau of Economic Research, Inc.

Incorporated under the Membership Corporation Laws of the State of New York, January 29, 1920

ITS ORGANIZATION AND PURPOSES

THE National Bureau of Economic Research was organized in 1920 in response to a growing demand for exact and impartial determinations of facts bearing on economic, social, and industrial problems.

It seeks not only to find facts and make them known, but to determine them in such manner and under such supervision as to make its findings carry conviction to Liberal and Conservative alike.

Entire control of the Bureau is vested in a Board of Directors, representing universities, learned and scientific societies, financial, industrial, agricultural, commercial, labor, and technical organizations.

Rigid provisions in the Charter and By-Laws guard the Bureau from becoming a source of profit to its members, directors, or officers and from becoming an agency for propaganda. No report of the Research Staff may be published without the approval of the Directors, and any Director who dissents from any finding approved by a majority of the Board may have such dissent published with the majority report.

The members of the Board of Directors are as follows:

THE SMOOTHING OF TIME SERIES

By

FREDERICK R. MACAULAY

OF THE STAFF OF THE

NATIONAL BUREAU OF ECONOMIC RESEARCH

INCORPORATED

NEW YORK

NATIONAL BUREAU OF ECONOMIC RESEARCH

INCORPORATED

1931

First printing, February, 1931

Printed in the United States of America by
J. J. LITTLE & IVES COMPANY, NEW YORK

RELATION OF THE DIRECTORS TO THE ECONOMIC WORK OF THE BUREAU

1—The object of the Bureau is to ascertain and to present to the public important economic facts and the interpretation thereof in a scientific and impartial manner, free from bias and propaganda. The Board of Directors is charged with the responsibility of ensuring and guaranteeing to the public that the work of the Bureau is carried out in strict conformity with this object.

2—The Directors shall appoint one or more directors of research chosen upon considerations of integrity, ability, character, and freedom from prejudice, who shall undertake to conduct economic researches in conformity with the principles of the Bureau.

3—The director or directors of research shall submit to the members of the Board, or to its executive committee when such is constituted and to which authority has been delegated by the Board, proposals in respect to researches to be instituted; and no research shall be instituted without the approval of the Board, or of its executive committee.

4—Following approval by the Board, or its executive committee, of a research proposed, the director or directors of research shall as soon as possible submit to the members of the Board, by written communication, a statement of the principles to be pursued in the study of the problem and the methods to be employed; and the director or directors of research shall not proceed to investigate, study, and report in detail, until the plan so outlined has been approved by the Board or the executive committee thereof.

5—Before the publication of the results of an inquiry the director or directors of research shall submit to the Board a synopsis of such results, drawing attention to the main conclusions reached, the major problems encountered and the solutions adopted, the nature of the sources from which the basic facts have been derived, and such other information as in their opinion shall have a material bearing on the validity of the conclusions and their suitability for publication in accordance with the principles of the Bureau.

6—A copy of any manuscript proposed to be published shall also be submitted to each member of the Board, and every member shall be entitled, if publication be approved, to have published also a memorandum of any dissent or reservation he may express, together with a brief statement of his reasons therefore, should he so desire. The publication of a volume does not, however, imply that each member of the Board of Directors has read the manuscript and passed upon its validity in every detail.

7—The results of any inquiry shall not be published except with the approval of at least a majority of the entire Board and a two-thirds majority of all those members of the Board who shall have voted on the proposal within the time fixed for the receipt of votes on the publication proposed; such limit shall be 45 days from the date of the submission of the synopsis and manuscript of the proposed publication, except that the Board may extend the limit in its absolute discretion, and shall upon the request of any member extend the limit for a period not exceeding 30 days.

8—A copy of this memorandum shall, unless otherwise determined by the Board, be printed in each copy of every work published by the Bureau.
Resolution adopted Oct. 25, 1926.

ACKNOWLEDGMENTS

The manuscript of this book has been read by Members of the Research Staff of the National Bureau and by the Bureau's Directors. I wish to acknowledge with thanks the helpful comments made by these gentlemen—Colonel M. C. Rorty in particular. I also desire to thank Dr. Henry Schultz, of the University of Chicago, who kindly read the manuscript. My obligations to Mr. Robert Henderson for specific suggestions are acknowledged in the text.

My sincere thanks are due to Miss Celeste Nason, who aided me in the earlier stages of the study, and to Miss Dorothy Achilles, my assistant in the later work. Miss Achilles not only made all the calculations and drew all the charts, but also read the manuscript with critical care.

F. R. M.

TABLE OF CONTENTS

APPENDICES

LIST OF CHARTS

THE SMOOTHING OF TIME SERIES

"Of Geology and Geognosy we know enough; what with the labours of our Werners and Huttons, what with the ardent genius of their disciples, it has come about that now to many a Royal Society, the Creation of a World is little more mysterious than the cooking of a dumpling; concerning which last, indeed, there have been minds to whom the question, *How the Apples were got in*, presented difficulties."

—*Sartor Resartus.*

INTRODUCTION

This little book grew out of an investigation into the history of interest rates and security prices in the United States since January, 1857. In handling the various series of monthly data prepared during that investigation, the problem arose of how comparisons between the series should be made. The relations between the different monthly series might have been analyzed by many statistical methods. We chose one of the simplest. We "smoothed" each data series in such a manner that relations between the larger [1] movements of the

[1] The object of smoothing the series included in the study of interest rates and security prices was to compare the "cyclical" movements of the various series. We wished to eliminate the multitude of minor movements, as we felt that their presence would obscure the picture of the major movements. For the discussion of the minor movements we relied primarily upon the raw data. However, for purposes other than ours, the investigator might wish to smooth his material in such a manner as to eliminate seasonal fluctuations and at the same time preserve not only the larger but most of the smaller movements of the data. The handling of such a problem is illustrated and discussed in Appendix I.

Any graduation of economic time series must, almost inevitably, be for a particular purpose only. The graduations presented in the study of interest rates and security prices are intended to supplement the data. They are not intended to replace the data. In this, they differ from adjustments made on physical observations in order to eliminate errors of measurement. They also differ from graduations which are intended to estimate the "universe" from a sample. The graduation of a mortality table is of this latter type.

various series became immediately apparent when the smooth curves were plotted on a chart. The particular method of smoothing or "graduation" which we used, eliminates from each series not only minor erratic movements but also monthly seasonal fluctuations. The resulting smooth curves will appear in the study of interest rates and security prices.

A brief chapter on the problem of smoothing was prepared for inclusion in that study. It covered little more than a description of the methods actually used. The manuscript was read by a few fellow statisticians. So many questions were then asked that the chapter was expanded. More questions were asked. Further expansion followed. What was originally designed as little more than a memorandum on a particular method grew into a rather general treatment of the whole problem of smoothing. It soon became apparent that a really simple treatment of even the elements of the subject required more space than a short chapter. As a long chapter on smoothing seemed somewhat of a digression, if included in an investigation of the history of interest rates and security prices, the National Bureau of Economic Research decided to publish this study separately.

An attempt has been made to present this intricate problem in as simple a manner as possible. This introduction is intended for the reader who wants merely a brief description of the nature of the smoothing process accompanied by some illustrations of a few simple graduations and directions for computing them. The more systematic discussion is contained in the body of the book.

A smooth curve may be described as one which does not change its slope in a sudden or erratic manner.[1] Smooth curves are not necessarily representable by simple mathematical equations. Indeed the expression, in its narrower sense, has often been reserved for curves which are not so representable. To suggest an adequate definition of smoothness is difficult, but it is still more difficult to suggest an adequate measure. The most commonly used mathematical measure of smoothness is based on the smallness of the sum of the squares of the third differences of successive points on the curve. This criterion amounts to measuring the smoothness of a curve by measuring how closely successive groups of four consecutive points can be described by second-degree parabolas.[2] Though such a concept may be useful, it certainly is not entirely logical. It implies that no curves are perfectly smooth except straight lines and second-degree parabolas.

[1] The mathematician may feel that this definition of smoothness ties up too definitely to mere second differences. However, do not the words "sudden or erratic manner" imply more than mere second differences?

[2] See note 1, page 54.

The fitting of a mathematical curve to physical observations may be a rational operation; the "smoothing" of economic time series is almost inevitably purely empirical. A mathematical curve fitted to observations made on bodies falling in a vacuum may be a statement of a law; the result of smoothing average monthly rates for Time Money on the New York Stock Exchange can hardly be more than a picture of what such rates would have been had they been unaffected by seasonal and erratic factors. It constitutes no law. However, in spite of the absence of any even hypothetically rational law, smoothing or "graduation" [1] seems useful for many purposes and therefore quite legitimate.

In the hands of a person who is thoroughly acquainted with all aspects of the data, freehand (or French curve) smoothing would seem to have much to recommend it. The most commonly advanced argument for freehand or graphic smoothing is that it saves time. As a matter of fact, one of the chief weaknesses of the method is the amount of time required—if the smoothing is to attain the object desired. If a curve be required which shall (1) be smooth, (2) give a good fit, and (3) eliminate seasonal fluctuations, the

[1] The terms *smoothing* and *graduation* are commonly thought of as not quite synonymous with *curve fitting*. The expression *curve fitting* should, perhaps, be reserved for the fitting to data of a curve representable by a mathematical equation.

amount of time used in correcting and recorrecting the freehand curve often becomes prohibitive. Moreover, most of the work must be done by the investigator himself. It is not the type of operation which can be delegated to a computer. Finally, the investigator himself may easily go astray. Judgment is, at best, a variable quantity. Unless the mathematical checks are so detailed that the freehand fitting practically amounts to a mathematical fitting by successive approximations, the investigator may easily describe the same data by distinctly different smooth curves if he does the fitting twice—with a month's time between operations. Most persons are incapable of good freehand smoothing. I do not hesitate to say, after having worked with many hundreds of students, that any fairly good mathematical method will, in at least nine cases out of ten, give better results than any method which requires much judgment.

Perhaps the best theoretical case for freehand smoothing can be made when there are reasons for suspecting that the underlying ideal curve is itself not smooth. If the underlying curve have cusps or be discontinuous, any continuous "smoothing" —whether mathematical or freehand—will, of course, somewhat obscure such characteristics. But the freehand method can easily smooth by parts—introducing any cusps or discontinuities which the investigator may wish in the "smooth" curve. For example, quotations for Call Money Rates or Bond Yields, before, during and after a financial panic, might be smoothed "up the hill" and "down the hill" but not over the top of the hill. Of course, one of the dangers in such a procedure, for purposes of comparison between various series, is that different draftsmen, or even the same draftsman at different times, will vary in their judgment as to when "over the top of the hill" should not be smoothed. Theoretically, freehand smoothing is ideal. Practically, it is a little like the faith of a mystic. It is conclusive evidence to the recipient of the vision alone.[1]

The simplest of purely mathematical methods of smoothing data is to take a moving average of the data and center that moving average. For example, a moving average, each value of which is the average of seven consecutive observations (which are equally spaced in time), may be used as a smoothed or theoretical value for the observa-

[1] Smoothing for purposes of comparing various series with one another is often useful even when the smoothing process has not described both series in a completely satisfactory manner. For example, the use of any time unit implicitly involves the use of one of the crudest and least adequate of mathematical smoothing processes—the simple moving average. To compare the annual production of pig iron during successive *years* with the annual volume of bank clearings during the same successive *years* is to compare selected points (twelve months apart) on the *12-months* moving average of the *monthly* production of pig iron with corresponding selected points on the *12-months* moving average of the *monthly* volume of bank clearings.

tion which is fourth in the list of the seven used in obtaining the particular moving average value.[1] Such a method of smoothing involves only extremely easy computation. It has, however, serious drawbacks. The resulting curve is seldom very smooth and it will not give a perfect fit to data except in ranges which can be adequately described by a straight line.[2] For example, a simple moving average, if applied to data whose underlying trend is of a second-degree parabolic type, falls always *within* instead of *on* the parabola. If applied to data whose underlying trend is of a sinusoidal type, it falls too low at maximum points and too high at minimum points. When applied to such data it cuts off tops and bottoms, usually resulting in a decidedly poor fit.

In general, if a type of smoothing be desired which shall, when applied to monthly data, eliminate seasonal and erratic fluctuations and at the same time give a smooth curve adequately describing the remaining cyclical and trend factors, something much more delicate than a simple 12-months

[1] If a 12-months moving average of monthly data be taken, any regular seasonal fluctuation in the data will, of course, be eliminated. Such a 12-months average should logically be centered between the sixth and seventh months. If a 2-months moving average of this 12-months moving average be taken, such average may be centered at the seventh month.

[2] Each point on a 12-months moving average, for example, is the middle point of a straight line fitted to 12 observations by the method of least squares.

moving average must be used. The 12-months moving average of any adequate graduation should approximate, or be a relatively good fit to, the 12-months moving average of the original data. In other words, the smooth curve should not itself be a 12-months moving average of the data, but a curve whose 12-months moving average is similar to the 12-months moving average of the data.[1]

Charts IV and V illustrate this characteristic. Chart IV shows a set of data (ninety-seven consecutive months of Call Money Rates on the New York Stock Exchange) and two smoothings or "graduations." The two graduations are (1) a 12-months moving average of the original data, and (2) a 43-term smooth curve fitted by a formula which we have used throughout our study of interest rates and security prices.[2] It is immediately

[1] Whenever the 12-months moving average of the data exactly equals the corresponding 12-months moving average of the smooth curve, the sum of 12 consecutive ordinates of the smooth curve, of course, exactly equals the sum of 12 consecutive ordinates of the data.

[2] This 43-term smooth curve is calculated as follows: Take a 5-months moving total of a 5-months moving total of an 8-months moving total of a 12-months moving total of the data. To the results apply the following simple weights: +7, −10, 0, 0, 0, 0, 0, 0, +10, 0, 0, 0, 0, 0, 0, −10, +7. Divide the final results by 9600. See pages 73, 74, 75.

Though the above procedure may seem complicated, it can easily be followed by any computer capable of taking a 12-months moving average. It takes about three and a half times as long to compute as a simple 12-months moving average.

apparent that the 43-term graduation is much *smoother* than the 12-months moving average. The difference in *goodness of fit* is shown in Chart V where a 12-months moving average of the data is compared with a 12-months moving average of each of the graduations. The reader will notice that the 12-months moving average of the 43-term graduation gives a much better fit to the 12-months moving average of the data than does the 12-months moving average of the 12-months moving average graduation. The 43-term graduation not only is much smoother but gives a much better fit.

As 43 months are needed to obtain one point on the 43-term graduation, there are necessarily 21 months at each end of the data which are not covered by the smooth curve, just as 6 months at each end of the data are not covered by the 12-months simple moving average. Such smoothings have to be extended, if they are to cover the entire range of the data. In the Call Money illustration above, this difficulty has been overcome by using data of the period before January 1886 and data of the period after January 1894. In the study of interest rates and security prices, each series was smoothed for the entire period January 1857 to date. Extension backwards in time was accomplished by using in each case the best data obtainable for the preceding 21 months. Forward in time, the graduations might have been mathematically extended without using any further data, real or hypothetical.[1] However, the easiest, and generally the best, way to extend a graduation forward in time is to apply it to hypothetical extrapolated data. This method has been used in the study of interest rates and security prices. The difficulties and dangers of such a procedure are distinctly less than the difficulties and dangers involved in freehand or other extrapolation of the smooth curves themselves. The tail end of any curve has necessarily a large probable error, and thoroughly adequate results—which would be likely to check with later data, when received—are generally quite improbable. This is just as true of graduations such as the Whittaker-Henderson, which need no extrapolation, as of graduations which require extrapolation. Moreover, mathematical extrapolation does not solve this difficulty.

The reader must not suppose that the particular 43-term formula emphasized in this book is presented as any final word on the subject of smoothing. It is primarily a method which is adapted to graduating monthly data in such a manner as to eliminate seasonal and erratic fluctuations and at the same time save all trend and the non-seasonal cyclical swings. It is not laborious. However, if the reader wishes to reduce the computation still

[1] For the details of the procedure to be used for such mathematical extension of the graduation, see pages 113, 114, 115.

further and yet obtain as good results as possible under such circumstances, he may use a simpler formula involving fewer steps. An example of such a formula would be: Take a 4-months moving total of a 7-months moving total of the data. Subtract a 16-months moving total of the data. Take a 3-months moving total of a 12-months moving total of the result. Divide by 432.[1] Twenty-nine observations are required to obtain one point on the graduated curve. The amount of computation is less than that involved in calculating the 43-term graduation. The formula will give comparatively good fits to a large range of sine curves.[2] It gives a comparatively good fit to our Call Money data, as may be seen by reference to column 14 of the table in Appendix VIII. It eliminates 12-months seasonal fluctuations. It is a fairly good formula for the investigator who wants a simple substitute for a 2-months moving average of a 12-months moving average. It takes little more than twice as long to compute as such a 2-months moving average of a 12-months moving average.

This 29-term formula falls an appreciable distance *outside* the parabola $y = x^2$, when applied

[1] This formula may also be applied as follows: Take a 14-months moving total of the data with the following simple weights: $-1, 0, 0, 0, +1, +1, +1, +1, +1, +1, 0, 0, 0, -1$. Take a 3-months moving total of a 3-months moving total of a 12-months moving total of the results. Divide by 432.

[2] See formula number 14 in Appendices IV, VII and VIII.

to points on that parabola.[1] If the investigator prefers a formula falling approximately on the parabola $y = x^2$, he may use the 29-term formula described on page 59. If he furthermore insists upon an absolutely irreducible minimum of labor, he must use a formula with a poorly shaped weight diagram. For example, a 27-term formula, which will eliminate monthly seasonal fluctuations, and give a distinctly better fit and smoother graduation than a 12-months moving average, but not as good a fit or smooth a graduation as can be obtained by using more complicated formulas, may be applied as follows: Take a 16-months moving total of the data with the following simple weights: $-1, 0, 0, 0, +1, +1, +1, +1, +1, +1, +1, +1, 0, 0, 0, -1$.[2] Take a 12-months moving total of this weighted 16-months moving total. Divide each of the final results by 72.[3] It is seldom advisable to use such an extremely simple formula.[4] A very little more labor will give distinctly better results.[5]

[1] This is usually an advantage with cyclical data.

[2] The reader will, of course, note that for calculation the middle set of units is treated as a simple 8-months moving total.

[3] If applied to points on the parabola $y = x^2$, the graduation will fall $\frac{1}{6}$ of a unit inside the parabola.

[4] Except in the case of the measurement of average seasonal fluctuations by means of operations on the deviations of the data from a graduated curve. For that purpose it is not necessary that the graduation be more than roughly adequate. See Appendix I.

[5] For example, the use of the following 27-term formula: Take a 10-months moving total of the data with the following simple weights: $-1, 0, 0, +1, +1, +1, +1, 0, 0, -1$. Take a 7-months

If the elimination of seasonal fluctuations is of very minor importance and if the short time fluctuations of the data are not too large (when compared with the amplitude of the longer time cyclical fluctuations), some of the well-known third-degree parabolic[1] summation formulas may be used without introducing any large element of erratic fluctuation.[2] For example, Kenchington's 27-term formula,[3] if applied to such a series as Railroad Stock Prices, gives fairly good results. If the investigator desires to follow his data somewhat more closely than Kenchington's formula permits, he may use such a formula as Spencer's

moving total of a 12-months moving total of the results. Divide by 168. The resulting graduation will be distinctly smoother and the fit better than with the slightly less laborious 27-term formula described in the text. The graduation falls $1\frac{5}{6}$ outside the parabola $y = x^2$.

[1] By a "third-degree parabola" we mean an equation of the form

$$y = A + Bx + Cx^2 + Dx^3$$

[2] Third-degree parabolic formulas always introduce some element of error in cyclical material. See pages 49, 50.

[3] A 5-months moving total of a 7-months moving total of an 11-months moving total of a 7-months moving total with the following simple set of weights: $-1, 0, +1, +1, +1, 0, -1$. Result divided by 385.

A 27-term formula which is easier to compute than Kenchington's and which gives appreciably closer fits to sine curves of short periods, with almost as close fits to sine curves of long periods, may be applied as follows: Take an 11-months moving total of the data with the following simple weights: $-1, 0, 0, +1, +1, +1, +1, +1, 0, 0, -1$. Take an 8-months moving total of a 10-months moving total of the results. Divide by 240. The weight diagram is only a little less smooth than Kenchington's. When applied to the parabola $y = x^2$ the formula gives a curve falling $\frac{1}{6}$ inside.

21-term formula.[1] However, the Spencer 21-term formula is very poorly adapted to smoothing such a series as Call Money Rates, where there is a large seasonal fluctuation and where the short-time fluctuations are large when compared with the cyclical fluctuations.

The advantages and disadvantages of the 43-term formula emphasized in this book and of a number of other methods of fitting are discussed in the text. For certain particular types of problem, the Whittaker-Henderson method of graduation, when judiciously used, is almost ideal. The Henderson method of computing this smooth curve is one of the most elegant contributions which have ever been made to the literature of the subject.

[1] See pages 51, 52, 53.

CHAPTER I

THE SMOOTHING OF ECONOMIC TIME SERIES. CURVE FITTING AND GRADUATION.

The statistical problem of fitting a mathematical curve to economic data has many of the characteristics of a problem in the adjustment of physical observations. Historically, the first problem of this kind arose in the adjustment of observations made on one variable. After a large number of observations had been made on one variable it was desired to obtain from those observations the most probable value of the variable, in other words, the most plausible estimate of the true value of the variable which could be made from the observations. The arithmetic mean of the observations was early proposed as a solution of this problem.[1]

The arithmetic mean was used before any attempt was made to develop a mathematical foundation for its use. An attempt at a mathematical foundation came with the theory of least squares. This theory claimed that the distribution of errors tends to be such that if a value of the variable be

[1] It is not necessary for this discussion to go into the question of when the arithmetic mean is the best value to use and when some other type of average would seem preferable.

taken which will make the sum of the squares of the deviations of the observations from this value a minimum, this value will be the most plausible estimate of the true value which could be made on the basis of the given observations. The arithmetic average of the observations is the value of the variable which makes the sum of the squares of the deviations a minimum.

The theory of least squares was extended to cover not only the problem of the most probable value of one variable but also the problem of the most probable relationship between two or more variables. If we think of measurements on one variable being charted as vertical deviations from a horizontal straight line, we see that the true value (which is free from errors of observation) may be represented by a horizontal straight line. We are attempting a numerical solution of the equation $Y = K$, where K is the unknown true value. When we approach the problem of the relations between two variables, we may think of our observations being charted with the values of one variable charted horizontally, and the values of the other variable charted vertically. If the relationship between the two variables is a straight line relationship, the problem is to find the best values for A and B in the equation $Y = A + BX$.

An example of a problem in two variables would be to discover a relation between altitude and barometric pressure. The investigator might first take a large number of observations on barometric pressure at different altitudes. For the purposes of this problem, let us assume that the altitudes have been accurately determined—in other words, let us assume that they are free from "error." Altitude is to be the "independent variable." The investigator may now draw points on a chart. Each point will represent a measurement of barometric pressure at a particular altitude. Altitudes are shown as horizontal distances measured from the left-hand edge of the chart and barometric pressures as vertical distances measured from the base of the chart. The problem is to draw a curve in among the points in such a manner that the ordinates (or vertical distances from the base) of the curve shall represent the most probable barometric pressures corresponding to the altitudes represented by the abscissas (horizontal distances from the left-hand edge). In other words, if we go along the base line 3 inches, which represents 3,000 feet altitude, the height of the curve at this point is to represent the most probable barometric pressure at 3,000 feet altitude.[1] The problem is not a problem of what is the most probable value of one variable

[1] Of course in the above problem, the barometric pressure readings are affected not only by *errors* but what is more important by actual variations in pressure which are not caused by altitude. We neglect the latter consideration in the illustration.

but what are the most probable values of one variable associated with definite values of the other variable. What is the most probable barometric pressure associated with any particular altitude?

In the solution of such a problem there are two parts. The first and generally the more important part is to decide what is the nature of the function relating the two variables to one another. The theory of least squares will give us little aid in making this decision. It may tell us which of two curves[1] gives the better fit; but it cannot draw our attention to such facts as that the curve giving the better fit may be grossly empirical, may not tie in with other known phenomena, and may become absurd immediately outside the range of the data, while the curve giving the poorer fit to the particular set of observations may be subject to no such theoretical drawbacks. The second part of the problem (if the curve be represented by an equation) is concerned with finding the "best" numerical values of the unknown constants. The theory of least squares states that these "best" numerical values of the constants will be such values as will make the sum of the squares of the vertical deviations of the observations from the fitted curve a minimum—in other words make the sum of the

[1] Such curves may or may not be simple mathematical equations covering the whole range of the data. For example, the laws of gases do not necessarily apply to liquids and solids.

squares of the deviations of the observational values from the corresponding theoretical values (on the curve) a minimum.[1]

Fitting a mathematical curve to economic time series is a statistical problem. It differs from an ordinary problem in the adjustment of observations in that the primary object of the fitting is not usually to eliminate errors of observation. The population of a country might conceivably be known exactly year by year for a long period and the statistician might wish to fit a curve to the whole series. Such a curve would not be fitted in order to discover what was the true population at any particular date. In other words, it would not be fitted to eliminate *errors*. It would probably be fitted for the entirely different purpose of discovering a law of population growth which would give a picture of what the population would have been if, throughout the period, it had been affected only by continuously acting and fundamental forces, excluding not only small and erratic influences but also unusual or disturbing influences, such as wars, famines, etc. Assuming some such law

[1] The method of least squares is, in practice, not so useful as the above might lead one to expect. Only if the equations relating the variables to one another are of a linear or parabolic type is the application of the method both direct and simple. With other types of equation, computation is, when not practically impossible, generally extremely laborious. Even the method of moments is not of universal application.

of population growth, the investigator might be interested in seeing how closely he could describe his particular population data by some simple growth curve, because he would be interested in speculating to what extent the continuously acting and typical forces were the controlling forces in the particular case he was examining.[1]

Trend is the outstanding characteristic of population curves. Seasonal fluctuations are negligible. Non-seasonal cyclical fluctuations are relatively small. On the other hand, there are time series in which little or no trend is apparent but which contain pronounced seasonal or cyclical movements. Monthly rainfall in the City of Seattle would be an example of a series having a pronounced "seasonal" fluctuation. Now the student of weather might wish to fit a curve to the monthly data for rainfall in the City of Seattle in such a manner as to eliminate seasonal and minor erratic fluctuations, but leave everything else. Even if the resulting curve showed no regular long time trend or regular and definite mathematical cycles, it might show considerably different amounts of rainfall at different periods.

There are economic time series which show both

[1] The investigator, of course, would not propose that his fitted population curve should, for all purposes, be substituted for the raw data. Graduations of, or curves fitted to, economic data can generally be used to replace the data for specific purposes only. For example, a graduation of monthly Call Money Rates has no

definite trend and pronounced and definite cycles. The production of eggs in the United States is an example of such a series. Annual production of eggs increases with the growth of the country, but there is, of course, a very pronounced 12-months or "seasonal" cycle. Of the series smoothed in the study of interest rates and security prices, Call Money Rates, Time Money Rates, Commercial Paper Rates and Railroad Bond Yields seem to be of a type showing no definite mathematical trend. At least, it would not seem feasible to propose a *law* of trend, such as has been proposed for the growth of population. The series for Railroad Stock Prices is of a type which shows evidences of a rather definite long time trend, though the cyclical fluctuations are extremely large. Bank Clearings and Pig Iron Production show smaller cyclical fluctuations and still more definite trend.

During most of the period covered by the interest rate study, Call Money Rates, Time Money Rates and Commercial Paper Rates show rather pronounced seasonal movements. Such seasonal movements as exist in the case of Bond Yields and Stock Prices are so small as to be practically negligible. It seems extremely improbable that any of the series contain definitely recurring periodic movements other than seasonal fluctuations.

such general replacement relation to the original data as a mortality curve has to its raw data.

The type of smooth curve which might be expected to appear in any particular time series if the series were unaffected by the minor or temporary factors which give rise to seasonal and erratic fluctuations is not necessarily representable throughout its length by any single simple mathematical equation. Its smoothness may be traceable to the fact that the ordinate of each point on the curve is correlated with the ordinate of the immediately preceding point. The curve may in this respect be similar to a curve resulting from the cumulation of a mere chance series. If twelve coins be thrown time after time and if the number of heads in each throw be charted as they occur, the resulting graph will not be a smooth curve nor will it tend to show any particular trend or cyclical fluctuations. Each throw stands by itself, uncorrelated with the preceding throws. However, if the series of throws are cumulated and the results charted, the graph will show considerable smoothness and both trend and cycles. The trend will most probably be representable by the straight line $y = 6x$. The cyclical appearance will be pronounced but the "cycles" will be irregular and not representable by any single simple mathematical equation. If, instead of cumulating the number of heads in each throw, the excess of the number of heads over 6 thrown each time be cumulated, the trend will vanish but the cycles will remain. The smoothness of such cumulative curves is inherent in their very nature. While not only the possible but the probable size of those larger movements which may be termed cyclical is greater in the cumulative than in the non-cumulative curve, both the possible and the probable size of those smaller movements which may be termed erratic is less. In the non-cumulative distribution of heads over 6 in number, the possible size of a "cyclical" movement is 12 (from −6 to +6). The possible "erratic movement" or change in size from one value to the next is also 12 (from +6 to −6 or from −6 to +6). If ten throws (12 coins thrown each time) be *cumulated* the possible size of a "cyclical" or major movement is increased from 12 to 60, while the possible "erratic movement" or *change* in size from one value to the next has decreased from 12 to 6 (+6 or −6). A similar though much less sensational story is told if *probable* rather than *possible* movements are considered. Cumulation smooths a series and introduces movements suggestive of trends and cycles.

Many economic time series seem to be of a type somewhat analogous to such cumulated chance series. Some economic series suggest chance series which have been cumulated twice. For example, each observation on a population series is not only highly correlated with the immediately preceding observation, but the first differences are highly correlated with the preceding first differences. The first differences are the resultant of three factors, births, deaths and migration. Births and deaths are functions of the size of the population and hence highly correlated with the same items for the preceding year. The excess of births over deaths and the amount of migration are correlated with similar items in the preceding year, though the correlation will not necessarily be so high as in the case of births or deaths alone. The commonest type of economic time series suggests a cumulated chance series on which has been superposed another but non-cumulated chance series and a more or less regular and unchanging seasonal fluctuation.

How should such series be smoothed? What sort of procedure would seem adapted to eliminate all seasonal and erratic fluctuations leaving a reasonable picture of the cyclical fluctuations and any underlying trend? While such a series as Railroad Stock Prices shows some evidence of an underlying trend which might be rationally described by some appropriate mathematical equation, its cyclical movements do not suggest any such possibility. Such series as Call Money Rates, Time Money Rates, Commercial Paper Rates and Railroad Bond Yields do not suggest any possibility of rationally describing either their long time trends or cyclical fluctuations by mathematical equations. To fit a straight line, a parabola, a compound interest curve, a logistic curve, or any other trend curve to such a series would seem highly empirical, not to say absurd.

Any attempt to fit a *single* mathematical equation to the whole of one of such series would seem quite as absurd if the equation were designed to describe periodic movements, as if it were designed to describe trend movements. The indiscriminate use of harmonic analysis to describe series which are not provably harmonic in their origin is indefensible. The position of each point on a fitted harmonic curve is affected by the position of each datum point. If short cycles appear in the first two thirds of the data, the harmonic curve will introduce short cycles in the last third even if they do not appear in the data. Any part of a smooth curve assumed to underlie the data curve is a smooth curve only because it is an outgrowth of the immediately preceding portion of such smooth curve. It is no more affected by distant points in the past than the wobbling track left by the rear wheel of a bicycle ridden by an inexperienced rider is affected by the position of the track 100 yards back.

Any part of the track of the rear wheel will be a comparatively smooth curve because of the manner in which it is made. It is an outgrowth of an immediately preceding portion of the track, but this does not mean that it is in any important sense affected by the track 100 yards back. A skillful rider could approach a large scale mathematical curve and ride upon it. Of course, he would not be able to do so instantly. He would have to come out of the particular curve on which he was traveling and gradually approach the mathematical curve on which he wished to travel. The attempt to represent the entire bicycle track by a single mathematical equation, whether designed to describe trend, "cyclical" movements or both, would be plainly quixotic. Harmonic analysis should not be used to describe data unless there are reasons for believing that cycles of constant period are inherent in the very nature of those data. Of course, the analysis itself may offer strong evidence that this is so, but evidently this could hardly be the case with any such bicycle "data." Attempts by various investigators to discover rigidly mathematical cycles (other than seasonal fluctuations) in economic data of the type presented in the study of interest rates and security prices have not so far been sufficiently convincing to make one feel that the above "bicycle" illustration is either far-fetched or illegitimate.

CHAPTER II

Graduating by Simple Moving Averages and by the Mid Ordinates of Third-Degree Parabolas Fitted by the Method of Least Squares.

If the data be such that it seems undesirable to fit any curve or graduation, the points of which will be appreciably [1] affected by the position of data points a long distance away from the graduated point, it would seem rather natural to consider using some form of "moving average." If the data are monthly and if the cycles are generally long and the erratic fluctuations are quite small, a simple 2-months moving average of a 12-months moving average [2] may give a comparatively smooth graduation which may be a fair approximation to the underlying curve desired. However, unless the cycles are long (72 months or more) such a simple moving average will not reach up into the maximum regions of the desired curve nor down into the minimum regions in the way that it should.

[1] Though each point on the Whittaker-Henderson curve (discussed below) is affected by the position of each point of the data, the influence of any distant data point is entirely negligible. This is unlike harmonic analysis where near and distant points have equal effects on the curve.

[2] Which eliminates seasonal fluctuations.

Moreover, unless the erratic fluctuations are quite small, a 2 of a 12-months moving average [1] will not contain enough terms to smooth the data at all successfully. Some longer moving average, such as an 8 of a 12, must then be used. Such a longer moving average will, however, require still longer cycles in the data if its "dampening" effect is not to be too great. For such reasons, moving averages are generally unsatisfactory.

In the preceding paragraph, we have referred to the fact that an 8-months moving average of a 12-months moving average will give a smoother curve than a 2 of a 12. The fact that the 8 of a 12 contains more terms than the 2 of a 12 is not the only reason why it gives a smoother result. Another reason is that its "weight diagram" is smoother. Before passing on to the discussion of types of smoothing other than simple moving averages, it seems desirable to explain the significance of smoothness in the weight diagram. A simple moving average may be thought of as being fitted to the data by the use of a set of weights whose total is unity. In a 13-months simple moving average, each weight would be $+1/13$. The weight diagram is a rectangle 13 units long and $1/13$ of a unit high. A 2-months moving average of a 12-months mov-

[1] That is, a 2-months moving average of a 12-months moving average. Similar abbreviations are used freely throughout the discussion which follows.

ing average also contains 13 weights, but the weight diagram is a little smoother. The end weights are each $+1/24$, while the other eleven weights are each $+1/12$. A commonly used criterion of smoothness is based on the sum of the squares of the third differences. Now the sum of the squares of the third differences of the 13-months moving average set of weights is $12/169$ while the sum of the squares of the third differences of the 2-months moving average of a 12-months moving average set of weights is only $1/72$.[1] A 4 of a 5 of a 6 contains only 13 terms but gives a still smoother weight diagram than a 2 of a 12. The sum of the squares of its third differences is only $1/800$.

A smooth weight diagram leads to smoothness in the resulting graduation because smoothing by means of any weighted or unweighted moving average amounts to distributing each observation over a region as long as the weight diagram and of the same shape as the weight diagram. For example, if we take a 13-months simple moving average of a series of observations, we may think of this 13-months moving average as being constructed as follows: Each individual observation is divided by 13 and these 13 equal fractional values

[1] In all discussions of weights (or weight diagrams) the reader must think of such weights as infinite in number; to the right and left of the actual values used, he must think of an infinite number of zero values. In connection with the discussion above, the reader may refer to the first few diagrams in Chart I.

are distributed along 13 adjacent points on the x axis, of which 6 are on each side of the observation.

A little thought or experimentation will quickly convince the reader that, so long as we restrict ourselves to *positive* weights, no moving average, weighted or unweighted, will exactly fit any mathematical curve except a straight line. If, to consecutive and equally spaced points on a second (or third) degree parabola, we fit a moving curve each point of which is the mid ordinate of a second-degree parabola fitted to n consecutive observations by the method of least squares, the fitted moving curve will, of course, fall on the original parabola.[1]

The mid ordinate of a second-degree parabola fitted by the method of least squares to n consecutive observations may be computed by means of a weighted moving average with a particular set of weights. For example, if n equals 13, the weights are:

$$-\frac{11}{143}, 0, +\frac{9}{143}, +\frac{16}{143}, +\frac{21}{143}, +\frac{24}{143}, +\frac{25}{143}, +\frac{24}{143}, +\frac{21}{143}, +\frac{16}{143}, +\frac{9}{143}, 0, -\frac{11}{143}.$$ [2]

[1] If fitted to a *third*-degree parabola, it will fall exactly on that curve, as the mid ordinate of a second-degree parabola fitted to data by the method of least squares is the same as the mid ordinate of a third-degree parabola fitted to the same data.

[2] Compare E. T. Whittaker and G. Robinson, *The Calculus of Observations*, Blackie & Son, 1924, p. 295. Also see Chart I.

This weighted average must, of course, be centered on the middle, or 7th, month. The points on such a graduation will naturally lie exactly on the original data points, as these data points are themselves points on a second (or third) degree parabola. We notice that there are two minus weights in the set of 13 weights. The graduation could not fall on the original parabola if it did not have these minus weights. A graduation of this type, each of whose points is the mid ordinate of a second-degree parabola fitted by the method of least squares, does not however necessarily give ideal results. There are at least four reasons for not using it to smooth such a time series as monthly Call Money Rates.

A first reason is that such a graduation will entirely eliminate seasonal fluctuations only by the most improbable accident. If, neglecting for the moment erratic fluctuations, the original monthly data be thought of as made up of two parts, (1) a smooth curve and (2) a seasonal fluctuation superposed on the smooth curve, the results of fitting a parabola to the original data are the same as if we fitted a parabola to the smooth curve and another parabola to the seasonal fluctuations and added together, each month, the pairs of resulting ordinates. Now, if the seasonal fluctuations were constant from year to year, the smooth curve fitted to them should, by the definition of seasonal fluctuations, be simply $y = 0$.[1] In general, a curve fitted to seasonals will give continuous zero values only if its weight diagram is such that equal weights are given to each *nominal* month. A simple 12-months moving average gives such equal weights to each nominal month. Many other graduations may be constructed which will do the same. Any moving average of a 12-months moving average will give equal weights to each nominal month. The 43-term graduation emphasized in this book gives such equal weights to each nominal month, but the weights for the mid ordinate of a second-degree parabola *fitted by the method of least squares* do not do so.

A second reason for not graduating such data as Call Money Rates by formulas for the mid ordinates of second-degree parabolas fitted by the method of least squares is that this method of graduation does not give the smoothest possible results. The weight diagrams are not even moderately smooth. For example, in the weight diagram

[1] The assumption is here made that in a seasonal movement the sum of the negative values in any 12-months period equals the sum of the positive values. If seasonal movement be defined in any other way, for example, as that the *product* of any 12 consecutive monthly seasonal values equals 1, the further assumption would have to be made that such a function of the data had been taken before smoothing as to permit the sum of the positive and negative values of the seasonal of this function to equal 0. For example, in the interest rate and security price series, this very product assumption is made and the 43-term curves are fitted not to the original data but to the logarithms of the original data. See Appendix VIII.

for the mid ordinate of a second-degree parabola fitted to 13 observations by the method of least squares, there are two distinct cusps at the $-{}^{11}/_{143}$ values. Unless a weight diagram be itself a smooth curve, it will not give the smoothest possible results in fitting.

A third reason for not using a graduation based upon the mid ordinates of second-degree parabolic curves is that such a graduation is poorly adapted to describing periodic functions. Second (or third) degree parabolic fitting may in exceptional cases be useful to describe long time trends but it is not adapted to the adequate description of cyclical or wave-like movements. If a curve based on moving parabolas be fitted to a sine curve by the method of least squares, the fitted curve will always be too low at maximum points and too high at minimum points. It is true that the smaller the number of terms to which such a parabola is fitted, the less will be the deviations of the fitted curve from the original sine curve. However, if we are not fitting to a sine curve but to actual irregular data, the taking of a small enough number of terms to allow the fitted parabola to reach up adequately into maximum sections of the data and down into minimum sections, will introduce the same difficulty encountered in the case of a simple moving average—though of course to a much smaller degree. Unless the erratic fluctuations of the data are very small as compared with the amplitude of the cyclical movements, a large number of terms will have to be used in the parabolic set of weights or the data will not be adequately "smoothed." However, unless the cycles of the original data have very long periods, it will not be possible to use a large number of terms without departing too far from the underlying fundamental curve.[1]

A fourth and very cogent reason for not smoothing by means of any graduation based on the mid ordinates of least squares parabolas is that the weights do not lend themselves to easy computation. For easy computation, the weights should always be such that they can be broken up into a short series of simple moving averages. If this cannot be done, we are faced with the multiplication of each datum point by its proper weight in order to get each separate point on the curve. If the number of weights in the weight diagram be large, the computation of the smooth curve then becomes extremely laborious.

[1] The negligible seasonal fluctuation in Railroad Bond Yields and the small size of the minor fluctuations as compared with the cyclical movements would have permitted the use of least squares second-degree parabolic formulas for smoothing that particular series.

CHAPTER III

Third-Degree Parabolic Graduation Formulas Having Smooth Weight Diagrams. Summation Formulas.

The actuaries have attacked the problem of obtaining sets of weights which lend themselves to easy computation and which, when fitted to second (or third) degree parabolas, will fall exactly on those parabolas but which, when fitted to irregular, non-mathematical data, will give smoother results than can be obtained from the weights which give the mid ordinate of a second-degree parabola fitted by the method of least squares.[1] A number of specialized formulas, each of which was designed to be applied to a specified number of terms of the data, have been developed by different investigators. Spencer's 21-term formula may be used as an illustration. The weights in that formula are:

$$-\frac{1}{350},\quad -\frac{3}{350},\quad -\frac{5}{350},\quad -\frac{5}{350},\quad -\frac{2}{350},\quad +\frac{6}{350},\quad +\frac{18}{350},$$

$$+\frac{33}{350},\quad +\frac{47}{350},\quad +\frac{57}{350},\quad +\frac{60}{350},\quad +\frac{57}{350},\quad +\frac{47}{350},\quad +\frac{33}{350},$$

$$+\frac{18}{350},\quad +\frac{6}{350},\quad -\frac{2}{350},\quad -\frac{5}{350},\quad -\frac{5}{350},\quad -\frac{3}{350},\quad -\frac{1}{350}.$$

[1] See Robert Henderson and H. N. Sheppard, *Graduation of Mortality and Other Tables*, the files of the Journal of the Institute of Actuaries, and the Transactions of The Actuarial Society of America,

Such a set of weights gives excellent results which are quite easy to compute. If the data were a monthly time series, the computer would first take a 7-months weighted moving total of the data with the following simple set of weights: -1, 0, $+1$, $+2$, $+1$, 0, -1. He then would take a 5-months moving total of a 5-months moving total of a 7-months moving total of the results. The final figures would then each be divided by 350.[1]

Such a formula as Spencer's 21-term formula tends to give a smooth graduation because the

[1] Note that $350 = 2 \times 7 \times 5 \times 5$. If superlative smoothness is not insisted upon, it is easy to discover a set of averages which will give a weight diagram such that when applied to a second (or third) degree parabola the results will fall on that parabola. The procedure may be illustrated from Spencer's 21-term formula above. Having picked the original set of weights, -1, 0, $+1$, $+2$, $+1$, 0, -1, on the basis of general judgment, the investigator calculates where a curve based upon such a set of weights would come if fitted to the second-degree parabola $y = x^2$. Multiplying the above set of weights by $+9$, $+4$, $+1$, 0, $+1$, $+4$, $+9$, he gets a total of -16 which divided by the total of the weights gives a figure of -8. In other words, this fitted curve will fall 8 units *outside* the parabola $y = x^2$. However, a simple 7-months moving average of the same parabolic data will fall 4 units *inside* the parabola and a simple 5-months moving average 2 units *inside* the parabola. If, after taking the original average, he continues by taking a 7-months moving average of the results and two 5-months moving averages, he raises the results of the first fitting by $+4$, $+2$, $+2$, or $+8$, bringing the fitted curve back on the parabola. A simple moving average of n consecutive terms comes $\frac{(n^2 - 1)}{12}$ C *inside* the parabola $y = A + Bx + Cx^2$.

A 21-term formula, which is often a desirable substitute for Spencer's, may be applied as follows: Take a 9-months weighted moving total of the data with the following simple set of weights:

weight diagram is itself a smooth curve. Recognition of the fact that the smoothness of the resulting graduation depends directly on the smoothness of the weight diagram led Robert Henderson to propose a solution of the problem of obtaining a general expression for the smoothest possible weight diagrams which, when fitted to second (or third) degree parabolas, would fall exactly on those parabolas. Assuming that lack of smoothness could be measured by calculating the sum of the squares of the third differences of the weights in the weight diagram, he developed a general formula which would make the sum of the squares of these third differences a minimum for any number

-1, 0, $+1$, $+1$, $+1$, $+1$, $+1$, 0, -1. Take a 3-months moving total of a 5-months moving total of a 7-months moving total of the results. Divide each of the final figures by 315.

The graduation resulting from the use of the above formula will tend to be a shade smoother than that resulting from the use of Spencer's 21-term formula. It will tend to fit cyclical data more adequately (it falls $\frac{2}{3}$ *outside* the parabola $y = x^2$). It is a shade simpler to compute.

Another 21-term formula which may be substituted for Spencer's is the following: Take a 9-months moving total of the data with the same weights as before, namely: -1, 0, $+1$, $+1$, $+1$, $+1$, $+1$, 0, -1. Take a 3-months moving total of a 4-months moving total of an 8-months moving total of the results. Divide each of the final figures by 288. This formula is almost exactly parabolic ($\frac{1}{6}$ outside $y = x^2$). It gives a weight diagram distinctly smoother than Spencer's 21-term formula. It fits short period sine curves somewhat less adequately than Spencer's formula. This is because it contains an 8-months moving average while Spencer's formula contains no average longer than 7. Like the preceding formula, it is a shade less laborious to apply than Spencer's 21-term formula.

of terms desired.[1] If the number of terms desired in a formula be represented by $2m - 3$, Robert Henderson shows that the general expression for the xth term is

$$\frac{315\{(m-1)^2-x^2\}\{m^2-x^2\}\{(m+1)^2-x^2\}\{(3m^2-16)-11x^2\}}{8m(m^2-1)(4m^2-1)(4m^2-9)(4m^2-25)}$$ [2]

To derive a set of 15 weights from the above general formula, 9 is substituted for the letter m and the values of the above expression are calculated for each value of x from -7 to $+7$. Making

[1] The successive orders of "differences" equal either the deviation or a multiple of the deviation of a datum point from a parabola put through as many consecutive data points as possible, excluding the particular datum point whose deviation is measured.

The successive differences are the *actual* deviations of a datum point from parabolas put through as many consecutive data points as possible which immediately *precede* the datum point under discussion. Thus, the nth difference is the deviation of a datum point from an $(n-1)$th degree parabola. For example, if the ordinates of five successive data points be 7, 17, 43, 91, 190, and if through the first four of these ordinates we put a third-degree parabola, its equation will be $y = 7 - 3x + 2x^2 + x^3$. Now the fifth ordinate given by this equation is 167. The deviation of the fifth data ordinate from the fifth ordinate of the curve is therefore $+23$. But this is the fourth difference of the five data ordinates.

The successive differences are *multiples* of the deviation of a datum point from a parabola fitted to other than immediately preceding data points. For example, let there be 8 consecutive data points. Through the first five data points and the last two data points put a sixth degree parabola. Then the deviation of the sixth datum point from this parabola will equal $\frac{1}{21}$ of the seventh difference of the eight data points. Notice that the seventh difference is $-y_1 + 7y_2 - 21y_3 + 35y_4 - 35y_5 + 21y_6 - 7y_7 + y_8$.

[2] Robert Henderson and H. N. Sheppard, *Graduation of Mortality and Other Tables*, p. 35; and Transactions of The Actuarial Society of America, Vol. XVII, p. 43.

the necessary substitutions in the formula, we obtain the following set of weights for a 15-term weight diagram:

$$-\frac{2652}{193154},\quad -\frac{3732}{193154},\quad -\frac{2730}{193154},\quad +\frac{3641}{193154},\quad +\frac{16016}{193154},$$
$$+\frac{28182}{193154},\quad +\frac{37422}{193154},\quad +\frac{40860}{193154},\quad +\frac{37422}{193154},\quad +\frac{28182}{193154},$$
$$+\frac{16016}{193154},\quad +\frac{3641}{193154},\quad -\frac{2730}{193154},\quad -\frac{3732}{193154},\quad -\frac{2652}{193154}.$$

The total, of course, equals unity.

The reader might be interested in comparing this "Ideal" formula with Spencer's 15-term formula.[1] Spencer's formula for 15 terms gives the following set of weights:

$$-\frac{3}{320},\quad -\frac{6}{320},\quad -\frac{5}{320},\quad +\frac{3}{320},\quad +\frac{21}{320},\quad +\frac{46}{320},\quad +\frac{67}{320},$$
$$+\frac{74}{320},\quad +\frac{67}{320},\quad +\frac{46}{320},\quad +\frac{21}{320},\quad +\frac{3}{320},\quad -\frac{5}{320},\quad -\frac{6}{320},$$
$$-\frac{3}{320}.$$

The sum of the squares of the third differences of these weights is only 12 per cent greater than the corresponding sum in Henderson's Ideal formula.[2]

[1] Spencer's 15-term formula is applied by taking a 5-months weighted moving total of the data with the following simple set of weights: -3, $+3$, $+4$, $+3$, -3. A 4-months moving total of a 4-months moving total of a 5-months moving total of the results are then taken. The final figures are each divided by 320. The actual computation of a Spencer 15-term graduation is given in Appendix II of this book.

[2] A 15-term strictly third-degree parabolic formula which is even better than Spencer's 15-term formula may be applied as fol-

This is an extremely low price to pay for the computation advantages of such a summation formula as Spencer's.[1]

Examining second-degree parabolic summation formulas and Henderson's Ideal formulas in the light of the four reasons already given for not using a graduation based on the mid ordinates of second-degree parabolas fitted by the method of least squares, we find that the first objection still holds. Such parabolic summation formulas do not eliminate seasonal fluctuation except by accident. However, they may be so constructed as to do so. All that is necessary to eliminate a 12-months sea-

lows. Take a 5-months weighted moving total of the data with the following simple set of weights: − 10, + 11, + 10, + 11, − 10. Take a 3-months moving total of a 4-months moving total of a 6-months moving total of the results. Divide the final figures by 864. The sum of the squares of the third differences is slightly less than in Spencer's formula. The shape of the weight diagram is distinctly closer to a Henderson's Ideal weight diagram. The formula is slightly less laborious to apply than Spencer's 15-term formula.

[1] The sum of the squares of the third differences of a set of weights which give the mid ordinate of a second-degree parabola fitted to 15 observations by the method of least squares is more than 26 times the corresponding sum in Henderson's Ideal formula. However, though this is bad, it is not quite as alarming as it sounds. The sum of the squares of the third differences of a graduation which has been applied to data by a least squares parabolic formula will not necessarily be 26 times the sum of the squares of the third differences of a graduation resulting from the use of the Henderson formula; in particular cases it may be very little greater. Unless many decimals are carried it will seldom be more than 3 or 4 times as great. Smoothness of the weight diagram, though important, is only one element in smoothness of the graduation.

sonal is to have a 12-months moving average as part of the calculation scheme. Henderson's Ideal formula cannot contain a 12-months moving average in its computation scheme as it is not computable by moving averages. It can therefore entirely eliminate seasonal fluctuation only by accident.[1]

The second objection to the use of weight diagrams based on the mid ordinates of second-degree parabolas fitted by the method of least squares is that such a method of fitting does not result in the smoothest possible curves. The third objection is based on the disadvantage of second (or third) degree parabolic fitting to periodic functions. Corresponding to these two objections is a sort of compound objection against the use of second-degree parabolic summation formulas or Henderson's Ideal formula. As parabolas are not periodic functions, if more than a small number of terms are used in a second-degree parabolic summation formula, the resulting graduation of any periodic type of data will tend to be unmistakably too low at maximum points and too high at minimum

[1] If a correct number of terms be used in a Henderson Ideal formula, it will practically eliminate seasonal fluctuations. For example, if the seasonal fluctuation be a 12-months sine curve, a Henderson 33-term Ideal formula will eliminate 101½ per cent of the seasonal. This little 1½ per cent over-elimination is, of course, generally negligible. A 25-term Ideal formula will eliminate only 76 per cent of such a sine seasonal, while a 37-term Ideal formula will eliminate 105 per cent of the seasonal.

points. However, if such a small number of terms be used in the formula as to overcome this difficulty, we are faced with the fact that in such series as monthly Call Money Rates the number of terms in the formula will not then be great enough to attain a high degree of smoothness.

Furthermore, it is not possible to have a smooth and well-shaped parabolic weight diagram which contains a 12-months moving average unless there are at least 27 terms in the diagram.[1] The second column of the table in Appendix III gives the weights for the smoothest possible 25-term second-degree parabolic graduation formula which will eliminate 12-months seasonal fluctuations. An examination of that table and Figures 7, 15, and 16 of Chart I will show the reader the inevitable hollow in the middle of the weight diagram, which comes in all 12-months weight diagrams which have an insufficient number of terms. The sum of the squares of the third differences of such an "ideal" 25-term, 12-months seasonal-eliminating set of weights is almost four times as great as in

[1] Even if such a summation formula as the Kenchington 27-term formula did not undesirably dampen the underlying cycles in such a series as Call Money Rates, it would not be usable if we insisted upon *rigid* elimination of 12-months seasonal fluctuations. If the seasonal fluctuation be a 12-months sine curve, Kenchington's formula will eliminate over 90 per cent of the seasonal. Spencer's 21-term formula will eliminate less than 45 per cent of such a seasonal. For Kenchington's 27-term formula, see note 3, page 29; for Spencer's 21-term formula, see pages 51, 52, 53.

the 25-term Henderson Ideal set of weights (which have no seasonal-eliminating condition).

For the investigator who wishes to use a 12-months seasonal-eliminating second-degree parabolic weight diagram (instead of such a diagram as Kenchington's 27-term formula, which does not entirely eliminate seasonal fluctuations), an excellent 29-term formula is the following: Take a 4-months moving total of an 8-months moving total of the data. Subtract a 17-months moving total of the data. Take a 2-months moving total of a 12-months moving total of the results. Divide by 360.[1] Such a formula, if fitted to the parabola $y = x^2$, gives results falling $\frac{1}{6}$ *outside* the parabola. The case would hardly ever occur in practice where this extremely small amount outside the parabola would not be an advantage rather than a disadvantage in fitting. The weight diagram is smooth and well shaped. The 29 weights are:

$$-\frac{1}{360},\ -\frac{3}{360},\ -\frac{5}{360},\ -\frac{6}{360},\ -\frac{5}{360},\ -\frac{2}{360},\ +\frac{3}{360},$$

$$+\frac{9}{360},\ +\frac{15}{360},\ +\frac{21}{360},\ +\frac{27}{360},\ +\frac{32}{360},\ +\frac{36}{360},\ +\frac{39}{360},$$

$$+\frac{40}{360},\ +\frac{39}{360},\ +\frac{36}{360},\ +\frac{32}{360},\ +\frac{27}{360},\ +\frac{21}{360},\ +\frac{15}{360},$$

[1] This formula may also be applied as follows: Take a 15-months moving total of the data with the following simple weights: -1, 0, 0, 0, $+1$, $+1$, $+1$, $+1$, $+1$, $+1$, $+1$, 0, 0, 0, -1. Take a 2 of a 3 of a 12-months moving total of the results. Divide by 360.

$+\frac{9}{360}, +\frac{3}{360}, -\frac{2}{360}, -\frac{5}{360}, -\frac{6}{360}, -\frac{5}{360}, -\frac{3}{360},$
$-\frac{1}{360}.$[1]

However, such a 29-term parabolic formula will not reach up to the maximum values of the underlying curve of such a series as monthly Call Money Rates, or down to the minimum values. Better results can often be obtained if we do not restrict ourselves to a curve which must exactly fit a second-degree parabola. For example, take such a formula as $\frac{1}{432}$ of a 3-months moving total of a 12-months moving total of the result of subtracting a 16-months moving total of the data from a 4-months moving total of a 7-months moving total of the data.[2] This formula contains 29 terms. That the weight diagram is not particularly close to the ideal form is not important. It falls $3\frac{1}{2}$ outside the second-degree parabola $y = x^2$. It is fitted to 97 months of Call Money Rates in column 14 of the table in Appendix VIII.[3]

[1] This 29-term approximately third-degree parabolic formula is used in the study of interest rates and security prices to obtain a graduation for *long time trend.* For that purpose the formula was applied to the Januaries, Mays and Septembers of the 43-term graduation. Intermediate values were read off a large scale chart on which a smooth curve had been drawn through these values by means of French curves. This procedure gave results quite accurate enough for our purposes and, of course, involved comparatively little work.

[2] See pages 26 and 27 and Figure 14, Chart I.

[3] See column 14 of the table in Appendix IV for the weights calculated to 5 decimals.

The fourth objection to the use of formulas giving the mid ordinates of second-degree parabolas fitted by the method of least squares is that the weight diagrams do not lend themselves to easy computation. Of course, this objection does not hold against the parabolic summation formulas but it holds in its fullness against Henderson's Ideal formulas.

Before concluding this chapter, it is desirable to draw the reader's attention to a relation between graduation by the mid ordinates of third-degree parabolas fitted by the method of least squares and graduation by third-degree parabolic formulas having smooth weight diagrams.

In graduation by the mid ordinates of third-degree parabolas fitted by the method of least squares, the assumption is made that the deviations of the data from the fitted parabola are given equal weights. On this assumption, the sum of the squares of the deviations is made a minimum. Graduation by third-degree parabolic formulas having smooth weight diagrams may be described as graduation by the mid ordinates of third-degree parabolas fitted by the method of least squares—when the various deviations of the data from the fitted parabola are given certain weights. The sum of the squares of the deviations of the data from the fitted parabola—when these deviations are given these specific weights—is made a minimum.

For example, if a third-degree parabola be fitted to 17 consecutive observations by the method of least squares, and if the seventeen deviations of the data from the curve be given the following positive weights— $\frac{1}{39}, \frac{1}{12}, \frac{2}{11}, \frac{1}{4}, \frac{1}{3}, \frac{5}{8}, \frac{6}{7}, 1, 1, 1, \frac{6}{7}, \frac{5}{8}, \frac{1}{3}, \frac{1}{4}, \frac{2}{11}, \frac{1}{12}, \frac{1}{39},$ the mid ordinate of the parabola fitted by the method of least squares will be the same value as obtained by applying Higham's[1] smoothing formula to the data.[2]

Graduation by Henderson's Ideal formula (see page 54) gives the same results as would be obtained by graduation by the mid ordinate of a third-degree parabola fitted to 2m — 3 observations by the method of least squares, if the weights, assigned to the successive deviations, *themselves* gave the smoothest[3] possible weight diagram. The formula for the relative weights assigned to the deviations in the case of Henderson's Ideal formula, is:

$$W_x = \{(m-1)^2 - x^2\}\{m^2 - x^2\}\{(m+1)^2 - x^2\}$$

The total of the 2m — 3 weights, of course, equals:

$$\frac{2}{35}(m^2-1)(4m^2-1)(4m^2-9)$$

[1] See note 1, page 133.

[2] Compare Robert Henderson's *Note on Graduation by Adjusted Average*. Transactions of the Actuarial Society of America, Vol. XVII, p. 45.

[3] Sum of the squares of the third differences a minimum.

CHAPTER IV

Fifth-Degree Parabolic Summation Formulas for Graduation.

If the investigator wishes both seasonal elimination and great power to smooth out erratic fluctuations in cyclical data without appreciably dampening the fluctuations of the underlying ideal curve, he must use something more flexible than a third-degree parabolic formula. He will naturally think of fifth-degree formulas. Approximately as close a fit to periodic data such as sine curves can be obtained with weights adapted to fit merely third-degree parabolas, only if the number of weights is radically reduced. However, less numerous weights tend to give a rougher graduation with non-mathematical data.

Sets of weights giving the mid ordinates of fourth (or fifth) [1] degree parabolas fitted by the method of least squares are not difficult to derive. A number are given in Whittaker and Robinson, *The Calculus of Observations*, page 296. The weight diagrams are themselves fourth-degree parabolas, just as the weight diagrams for the mid

[1] Of course, any symmetrical weight diagram which will fit a fourth-degree parabola will also fit a fifth-degree parabola.

ordinates of second (or third) degree parabolas are themselves second-degree parabolas. For example, the fourth (or fifth) degree weights for 13 observations are:

$$+\frac{110}{2431}, \quad -\frac{198}{2431}, \quad -\frac{135}{2431}, \quad +\frac{110}{2431}, \quad +\frac{390}{2431}, \quad +\frac{600}{2431},$$
$$+\frac{677}{2431}, \quad +\frac{600}{2431}, \quad +\frac{390}{2431}, \quad +\frac{110}{2431}, \quad -\frac{135}{2431}, \quad -\frac{198}{2431},$$
$$+\frac{110}{2431}.$$

Cyclical data or periodic functions can, of course, be more closely approximated by such fifth-degree parabolic formulas than by corresponding third-degree formulas. More observations can therefore be used. This leads to the possibility of greater smoothness in the graduation. However, the method is open to the same criticisms as third-degree least squares graduation in that it does not eliminate seasonal fluctuations and the fitted curve is extremely laborious to compute.

Fifth-degree parabolic summation formulas with smooth weight diagrams may be constructed almost as readily as third-degree formulas. Such summation formulas are free from a number of the disadvantages of fifth-degree least squares formulas. The amount of computation is much less. The smoothness of the results obtained by fitting to actual data is appreciably greater. Seasonal fluctuations are eliminated if the formula be properly constructed. For example, a smooth weight diagram which, if applied to a fifth-degree or any lower order of parabola, will fall exactly on that parabola and which will eliminate 12-months seasonal fluctuations, may be computed as follows: Take successively a 2-months moving total of the data, then a 3-months moving total, then another 3, then a 4, a 6, an 8, a 10, and a 12-months moving total. To the results apply the following set of weights: + 24,374, − 100,301, + 152,034, − 100,301, + 24,374. Divide each of the final results by 74,649,600.[1] The 45 weights, to the nearest fifth decimal, are given in column 22 of the table in Appendix IV. Figure 22 of Chart I gives a picture of the weight diagram.

With such a set of fifth-degree parabolic weights it is, of course, possible to obtain a closer fit to data than with a set having the same number of weights but designed to fit merely second (or third) degree parabolas. This particular 45-term fifth-degree parabolic set of weights, if applied to monthly Call Money Rates, gives results resembling those obtained by means of the 43-term curves we have used in the study of interest rates and security prices. The amount of computation

[1] In actual computation this final step is unnecessary. The 5 weights are each divided by 74,649,600 and the results (taken to a sufficient number of decimals) are used instead of the 5 weights above.

is, however, much greater than that required for the 43-term formula.

The reader must remember that, though each of two weight systems contain 45 terms and each be designed to fall on fifth-degree and all lower order of parabolas, it is not necessary that they show the same fit when applied to other curves than parabolas. A 45-term fifth-degree parabolic summation formula which contains fewer moving totals and is therefore easier to compute than the one which has just been described and which also gives a closer fit to sine curves, may be applied as follows: Take successively a 3-months moving total of the data, a 5-months moving total, another 5, an 8, and a 12-months moving total. To the results apply the following set of weights: + 1,331,771, − 1,949,056, 0, 0, 0, 0, 0, 0, + 2,175,370, 0, 0, 0, 0, 0, 0, − 1,949,056, + 1,331,771. Divide each of the final results by 6,773,760,000.[1]

Fifth-degree parabolic weight systems containing a smaller number of terms can, of course, be constructed. The resulting smooth curves will generally follow the data more closely but will also tend to be somewhat less smooth and more affected by merely accidental irregularities and non-typical movements in the data. A smooth 43-term formula which will eliminate 12-months seasonal fluctua-

[1] See note 1, page 65, and Column 21 of the table in Appendix IV and Figure 21, Chart I.

tions and which, if fitted to a fifth degree or lower order of parabola will fall exactly on that parabola, may be applied as follows: Take successively a 5-months moving total of the data, then another 5, an 8, and a 12. To the results apply the following set of weights: + 3,819,893, − 5,598,848, 0, 0, 0, 0, 0, 0, + 6,380,310, 0, 0, 0, 0, 0, 0, − 5,598,848, + 3,819,893. Divide each of the final results by 6,773,760,000.[1]

In a similar manner a 41-term fifth-degree parabolic set of weights, which will eliminate 12-months seasonal fluctuations, may be applied as follows: Take successively a 3-months moving total, a 5, an 8, and a 12-months moving total. To the results apply the following set of weights: + 1,158,703, − 1,703,808, 0, 0, 0, 0, 0, 0, + 2,031,-010, 0, 0, 0, 0, 0, 0, − 1,703,808, + 1,158,703. Divide each of the final results by 1,354,752,000.[2]

A 35-term strictly fifth-degree parabolic set of weights, which will eliminate 12-months seasonal fluctuations, but which has a weight diagram that is only moderately smooth and well shaped, may be applied as follows: Take successively 3, 5, 8 and 12-months moving totals of the data. To the results apply the following set of weights: + 273,632, − 472,175, 0, 0, 0, + 469,086, 0, 0, 0,

[1] See note 1, page 65, and Column 20 of the table in Appendix IV and Figure 20, Chart I.

[2] See note 1, page 65, and Column 19 of the table in Appendix IV and Figure 19, Chart I.

− 472,175, + 273,632. Divide each of the final results by 103,680,000.[1]

Similar strictly fifth-degree formulas may be constructed containing even fewer weights than 35. A 33-term formula which is not strictly fifth-degree (though nearly so) but which is less laborious to use than the preceding strictly parabolic formulas may be applied as follows: Take successively 5, 8 and 12-months moving totals of the data. To the results apply the following simple weights: + 58, − 100, 0, 0, 0, + 100, 0, 0, 0, − 100, + 58. Divide each of the final results by 7680.

The fact that such a formula will not exactly fit a fifth-degree parabola is of negligible significance.[2] When used for graduating such a series as the 97-months of Call Money Rates printed in Appendix VIII, it gives results differing little from those obtained by using the 35-term strictly fifth-degree parabolic formula already described. The graduation is a shade less smooth than that given by the 35-term formula. The formula fits sine curves of short period a shade closer than the 35-term formula. It does not fit those of long period quite so well.

When the number of weights in a fifth-degree

[1] See note 1, page 65, and Column 18 of the Table in Appendix IV and Figure 18, Chart I.

[2] See pages 70, 71.

parabolic, 12-months seasonal eliminating formula is reduced much below 33 or 31, the resulting graduation of such data as early monthly Call Money Rates tends to show a large increase in the number of points of inflection and even in the number of maxima and minima. The graduated curve tends to "weave" among the data. Artificial and arbitrary sinuosities make their appearance. Such peculiar behavior is not difficult to explain. When the number of weights is reduced much below 33 or 31, it becomes impossible to obtain a smooth and well-shaped weight diagram.[1] This impossibility is the result of insisting upon sea-sonal eliminating qualities. The reader will understand how great a restriction this imposes on the shape of the weight diagram when he remembers that the formula is necessarily so constructed

[1] The smallest possible number of weights which will eliminate 12-months seasonal fluctuations and fall on fifth and all lower orders of parabolas is sixteen. However, the points on the smooth curve resulting from the use of such a weight diagram would have to be centered between data dates. With a 17-weight diagram the points on the curve will be centered on data dates. Such a 17-weight diagram necessarily involves the following operations and no others: Take a 12-months moving total of the data. To the results apply the following set of weights:

$$+\frac{3775}{3456},\ -\frac{12201}{3456},\ +\frac{8570}{3456},\ +\frac{8570}{3456},\ -\frac{12201}{3456},\ +\frac{3775}{3456}.$$

There are 17 weights in all. To the first two decimals the weights are: +1.09, −2.43, +.04, +2.52, −1.00, +.08, +.08, +.08, +.08, +.08, +.08, +.08, −1.00, +2.52, +.04, −2.43, +1.09: Total = +1.00. This gives a decidedly bizarre weight diagram, as may be seen from examining the figures.

that it gives the same graduation if applied to the raw data as it does if applied to the data after an adjustment for seasonal fluctuations—*no matter what the seasonal fluctuations may be assumed to be.* Now, if 33 or more terms are taken in the formula, the seasonal fluctuations, *however absurd they may be assumed to be*, are taken care of quite easily by the mere fact of their repetition. However, as the number of terms is reduced, a point is reached at which the repetition of the seasonal fluctuations is not great enough. The formula can then no longer handle the data adequately. The danger line seems to be reached about where it becomes necessary to use a formula having two distinct modes instead of merely one.[1]

While good results can be obtained by fitting strictly fifth-degree parabolic formulas to series such as Call Money Rates, we must not forget that parabolic formulas are purely empirical. There is no logical reason why we should insist upon a formula which will exactly fit a fifth-degree parabola. For most time series an even more important consideration than how closely the formula will fit a fifth-degree parabola is how closely it will fit sine curves of various periods. Moreover, if we do not insist upon an exact fit to a fifth-degree parabola, it is possible easily to develop formulas which

[1] For an example of a third-degree parabolic duo-model formula, see Appendix III.

will give *approximate* fits to such fifth-degree parabolas and also to sine curves of a large range of period, and at the same time be extremely simple to compute. As an example of such a formula, take a 3-months moving total of a 5-months moving total of an 8-months moving total of a 12-months moving total of the data. To the results apply the following extremely simple set of weights: +2, −3, 0, 0, 0, 0, 0, +3, 0, 0, 0, 0, 0, −3, +2. Divide the final results by 1440.[1]

If the 39-term weight system resulting from the above simple procedure be applied to 39 terms of the parabola $y = x^2$ (from $x = -19$ to $x = +19$), the resulting point on the fitted curve will be $-1/6$, in other words the fitted curve will, at this point, fall $1/6$ outside the parabola. If the formula be similarly applied to the parabola $y = x^4$, the resulting point will be $+415\,3/10$. In other words the fitted curve at this point will fall a little more than 415 inside the parabola $y = x^4$. However, as the second-degree parabola above has a range from zero to +361, and the fourth-degree parabola a range from zero to +130,321, the 39-term formula accounts for 100.05 per cent of the term involving the second power of x and 99.68 per cent of the

[1] Actual computation is simplified by using the weights: +10, −15, 0, 0, 0, 0, 0, +15, 0, 0, 0, 0, 0, −15, +10, and dividing by 7,200.

For weights and weight diagram, see Column 23 of the table in Appendix IV and Figure 23, Chart I.

term involving the fourth power of x in the equation $y = A + Bx + Cx^2 + Dx^3 + Ex^4 + Fx^5$.[1] The formula, of course, fits exactly all the other terms in the equation. This 39-term formula must therefore give an extremely close approximation to any fifth-degree parabola to which it may be fitted. An examination of column 23 of the table in Appendix VII will give the reader an idea of how adequate a fit such a formula will give to a large range of sine curves. From that table, it may be seen that it gives an adequate fit to a larger range of sine curves than any one of the five strictly fifth-degree parabolic formulas described above. Even the 35-term fifth-degree formula gives a closer fit only to sine curves of a shorter period than 19 months,[2] in spite of the fact that the 39-term formula contains enough terms to free it almost entirely from most of the undesirable sinuosities which show a tendency to appear when strictly parabolic formu-

[1] The 33-term formula accounts for 99.84 per cent of the term involving the second power of x and 99.90 per cent of the term involving the fourth power of x.

[2] When fitted to sine curves whose periods are around 40 months, both this 39-term approximately parabolic formula and the 43-term formula discussed immediately below give results falling a fraction of 1 per cent of the amplitude of the sine curve *outside* the sine curve. This is, of course, entirely negligible. It could have been corrected by a slight change in the moving totals used in computation, but such a change does not seem worth making, as any correction of this negligible amount outside the sine curve in the 40 months period range would tend to give an appreciably poorer fit to sine curves of shorter period than 30 months.

las of few terms are used. The results of applying this 39-term formula to 97 months of Call Money Rates may be seen from examining Column 23 of the table in Appendix VIII.

A somewhat similar approximately fifth-degree formula, containing 43 terms, may be applied as follows: Take a 5-months moving total of a 5-months moving total of an 8-months moving total of a 12-months moving total of the data. To the results apply the following extremely simple weights: + 7, − 10, 0, 0, 0, 0, 0, 0, + 10, 0, 0, 0, 0, 0, 0, − 10, + 7. Divide the final results by 9600. There are 43 weights in all.[1]

When such a 43-term formula is applied to actual data, it gives a curve which is slightly smoother and freer from any hint of sinuosity than even the 39-term formula discussed immediately above. The results of applying such a 43-term formula to sine curves of various periods may be seen by examining Column 24 of the table in Appendix VII. The results of applying it to 97 months of Call Money Rates may be seen from the table in Appendix VIII. Like the 39-term formula, this 43-term formula is approximately parabolic. It corrects for 99.96 per cent of the term involving the second power of x and 99.59 per cent of the term involving the fourth power of x in the equa-

[1] For the weights, see Column 24 of the table in Appendix IV. For the weight diagram, see Figure 24, Chart I.

tion $y = A + Bx + Cx^2 + Dx^3 + Ex^4 + Fx^5$, and, of course, 100 per cent of all the other terms. It is the curve we have used throughout the study of interest rates and security prices.

The fact that neither the above 39-term formula nor the above 43-term formula will give a graduation falling exactly on a fifth-degree parabolic curve if fitted to such a curve is a matter of no practical significance. Even if the graduation deviated considerably further from such a fifth-degree parabolic curve than it actually does, such deviation would not necessarily be important. There is nothing magical in parabolic curves. It is true that, if enough terms be taken in a parabolic or power expansion, it may be made to represent almost any curve whatever. However, such representation is purely empirical and, if too many terms be taken in the parabolic expansion, leads to impossible and ridiculous results immediately outside the range of the data.[1]

For most economic time series, approximation to any particular section of the curve is more naturally accomplished by means of sine curves than by means of parabolas. The reader must remember that such series are commonly of an un-

[1] Similar purely empirical fitting may be performed by means of harmonic analysis. In this case, the particular curve is not approximated by means of successive parabolas, but by means of a series of sine curves.

mistakably cyclical type. For smoothing series such as those handled in our study of interest rates and security prices, both the 39-term approximately parabolic formula and the 43-term approximately parabolic formula are to be highly recommended.[1]

Before leaving this section of the discussion, the reader might care to examine Appendix IV, in which the weights are given for each of the formulas just discussed,[2] Chart I in which the weight diagrams are presented, Appendix VII in which the results of applying the various formulas [2] to sine curves of different periods are tabulated, and Appendix VIII, in which are given the graduations which result from applying the various formulas [2] to 97 months of Call Money Rates.

The figures in Chart I are weight diagrams corresponding to the following weight systems:

Fig. 1. A 12-months simple moving average.
Fig. 2. A 2-months moving average of a 12-months moving average.
Fig. 3. An 8-months moving average of a 12-months moving average.
Fig. 4. A 4-months moving average of a 5-months moving average of a 6-months moving average.

[1] Even the 33-term formula may be used with confidence, and it involves one less computation step than either the 39-term or the 43-term. On the other hand, there are some slight advantages in using a formula with more than 33 terms. The graduation is smoother and there is less chance of its ever being more than microscopically affected by the assumption of constant seasonal fluctuations over only a short period.

[2] Except the 33-term formula.

CHART I

GRADUATION WEIGHT DIAGRAMS

A GRAPHIC REPRESENTATION OF THE WEIGHTS IMPLIED IN TWENTY-FOUR SMOOTHING FORMULAS

IN EACH DIAGRAM
THE SUM OF THE ORDINATES EQUALS UNITY

FIG. 1

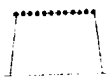


FIG. 2

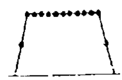


FIG. 3

FIG. 4

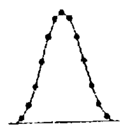


FIG. 5

FIG. 6

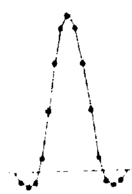


FIG. 7

FIG. 8

CHART I

GRADUATION WEIGHT DIAGRAMS
(CONTINUED)

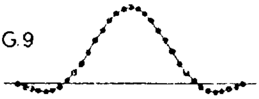


FIG. 9

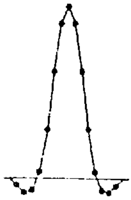


FIG. 10

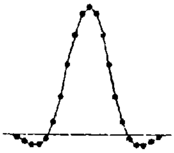


FIG. 11

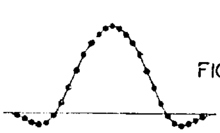


FIG. 12

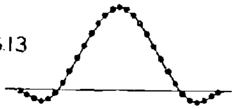


FIG. 13

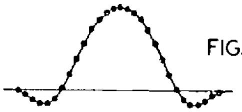


FIG. 14

FIG. 15

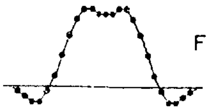


FIG. 16

CHART I

GRADUATION WEIGHT DIAGRAMS
(CONCLUDED)

FIG. 17

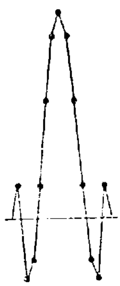


FIG. 18

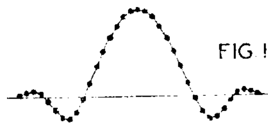


FIG. 19

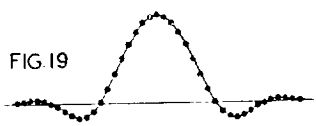


FIG. 20

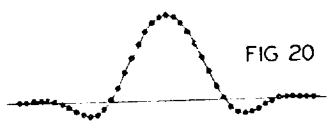


FIG. 21

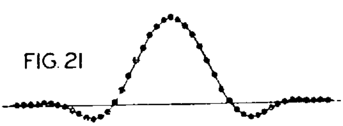


FIG. 22

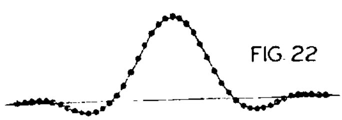


FIG. 23

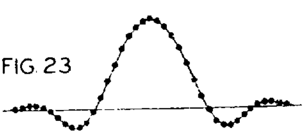
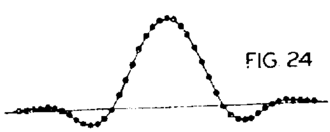


FIG. 24



When examining the preceding weight diagrams, the reader must remember that the weights are infinite in number. To the right and left of the weights represented by black dots, he must think of an infinite number of zero weights (lying on the horizontal straight line). For example, Figure 5 must be considered an extremely rough weight diagram, in spite of the fact that all the real plus or minus weights lie on a second-degree parabola. The *complete* weight diagram (which includes the zero weights) shows four violently disturbing angles—two at the last real weights (represented by black dots) and two at the first zero weights (on the straight line).

CHAPTER V

Dr. E. C. Rhodes' Graduation Proposal

In 1921, Dr. E. C. Rhodes proposed a method of smoothing based on the fitting of fourth-degree parabolas.[1] The method is somewhat different from any other discussed in this book.

The graduation is accomplished as follows: First, from an examination of the data, decide upon a certain number of terms—for example, 15. Then, to the first 15 observations fit a fourth-degree parabola by the method of least squares. Take the first 8 points on this parabola as the first 8 points of the desired smooth curve. To the second 15 observations (points 2 to 16, inclusive) fit a fourth-degree parabola by the method of least squares in such a manner that this second parabola passes through the 8th point of the first parabola and at that 8th point has the same slope as the first parabola. Obtain the 9th point on the smooth curve from this second parabola. Repeat the operation for each successive point on the smooth curve until such a "conditioned" parabola has been fitted to the last 15 points of the data. Use the

[1] E. C. Rhodes, *Smoothing*—Tracts for Computers, No. 6. Edited by Karl Pearson, Cambridge University Press, 1921.

last 8 points on this parabola as the last 8 points of the smooth curve.

This method of fitting can evidently give some strange results, as may be seen from an examination of the tables and charts in Dr. Rhodes' monograph. The closeness of fit of the graduation to the test data [1] is not connected in any simple manner with the number of terms to which the fourth-degree parabola is fitted. Decreasing the number of terms sometimes increases and sometimes decreases the goodness of fit. For example, Dr. Rhodes presents four graduations which he designates a 9-point curve, an 11-point curve, a 13-point curve, and a 15-point curve. A priori, one might think that, though the 9-point curve might not be as smooth as the 13 or the 15, it would lie closer to the data. Now, the sum of the squares of the deviations of the data from the fitted curve is much greater in the case of the 9-point curve than in the case of either the 13 or the 15. For the 29 observations which are covered by all four curves and which he uses as a comparison range, the sum of the squares of the deviations of the data from Dr. Rhodes' smooth curves are: for the

[1] Dr. Rhodes' test data are taken from an article by W. F. Sheppard entitled *Graduation by Reduction of Mean Square of Error*, Journal of the Institute of Actuaries, Vol. 48, p. 178. The data are rates of infantile mortality from causes other than diarrheal diseases, for the 42 years from 1870 to 1911, inclusive.

15-point curve, 266.0; for the 13-point curve, 262.2; for the 11-point curve, 243.6; and for the 9-point curve, 348.7.[1] In Dr. Rhodes' illustration, the 9-point curve is weirdly erratic, as may be seen from a glance at Diagram II at the end of his monograph.

Dr. Rhodes' fitted curves are smooth when the correct number of data items are taken in fitting his parabolas; otherwise strange, sweeping sinuosities are introduced which have little apparent relation to the data. For the 42 observations, the sum of the squares of the third differences of Rhodes' 9-point curve is 98.52. This compares with 11.75 for the 11-point curve, 5.58 for the 13-point curve and 1.61 for the 15-point curve. Dr. Rhodes was not pleased with the results he obtained by fitting the 9-point and 11-point curves. An examination of the table in Appendix V of this book, or Diagram II in the Appendix to Dr. Rhodes' book, will suggest the reason why. The sinuosities introduced have little apparent relation to the data. Dr. Rhodes was pleased with the results of his 13- and 15-point curves.

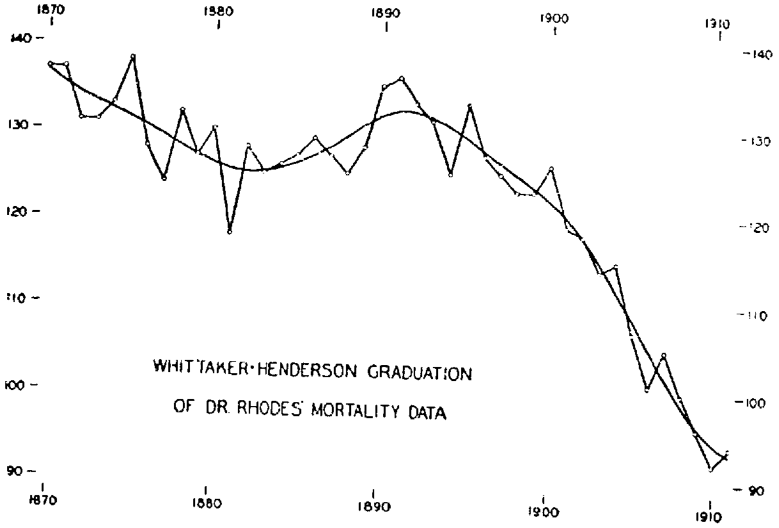
In Appendix V are given: I, Rhodes' data; II, Rhodes' 15-point curve; III, Rhodes' 13-point curve; IV, Rhodes' 11-point curve; V, Rhodes' 9-point curve; VI, Sheppard's graduation; VII,

[1] Rhodes, *Smoothing*, p. 25.

Spencer's 15-term graduation; VIII, a Whittaker-Henderson graduation with n = 2.[1]

Though the figures for the Spencer 15-term curve and the Whittaker-Henderson curve are given to two decimals in Appendix V, all the comparisons which follow are from calculations based on nearest first decimal places, as Dr. Rhodes' figures con-

CHART II



tain only one decimal. The first six columns of the table in Appendix V are from Dr. Rhodes' monograph. Columns VII and VIII were calculated in the National Bureau of Economic Research. Sheppard's graduation is not particularly interesting, though it seems to have been the point of departure for Dr. Rhodes' monograph. Spencer's

[1] For the Whittaker-Henderson method of graduation see Appendix VI.

15-term formula is attractive. The amount of labor involved is, even if the ends be extrapolated, only a small percentage of that necessary in the case of the Rhodes' method. The Whittaker-Henderson graduation gives the best results.

Some comparisons of the Whittaker-Henderson graduation and the Spencer 15-term graduation with the various results obtained by Dr. Rhodes' method may be interesting. The table below gives comparisons based upon the middle twenty-eight observations and the total forty-two observations. This classification is introduced in order to give a comparison of the different curves in the middle region where no extrapolation occurs, as well as for the whole range. The table consists of three parts. Part I compares the lack of smoothness of the various curves; Parts II and III give measures of their badness of fit.

MEASUREMENTS ON THE VARIOUS GRADUATIONS OF DR. RHODES' TEST DATA.

Part I.

Sums of the Squares of the Third Differences of the Various Curves.

	28 Observations	42 Observations
Rhodes' 15-point curve	1.10	1.61
" 13 " "	3.55	5.58
" 11 " "	7.95	11.75
" 9 " "	55.67	98.52
Dr. Sheppard's Graduation	39.45	45.12
Spencer's 15-term formula	1.46	2.42
Whittaker-Henderson ($n = 2$)	.77	1.15

PART II.

Sums of Deviations of Data from Curves.

(without regard to sign)

	28 Observations	42 Observations
Rhodes' 15-point curve	65.1	96.2
" 13 " "	65.9	97.3
" 11 " "	61.4	91.8
" 9 " "	69.9	109.7
Dr. Sheppard's Graduation............	68.1	99.7
Spencer's 15-term formula	64.0	93.9
Whittaker-Henderson (n = 2)	66.2	96.3

PART III.

Sums of the Squares of the Deviations of the Data from the Curves.

	28 Observations	42 Observations
Rhodes' 15-point curve	261.59	368.42
" 13 " "	257.77	361.67
" 11 " "	239.24	337.74
" 9 " "	345.07	488.85
Dr. Sheppard's Graduation	259.95	370.55
Spencer's 15-term formula	250.56	353.51
Whittaker-Henderson (n = 2)	256.70	359.17

There seems no good reason for substituting such a method as that of Dr. Rhodes for the much simpler summation method illustrated by Spencer's 15-term formula. While the results obtainable by the Whittaker-Henderson method are distinctly better than those obtainable by any one of the other methods, Dr. Rhodes cannot be criticized for neglecting this method, as it had not been published at the time he wrote his monograph.

Dr. Rhodes stresses the fact that his curve is a continuous curve,[1] each part of which may be represented by a mathematical equation, and that therefore interpolation or summation may be exactly accomplished. This is not highly important. Not only the Whittaker-Henderson curve, but also most good summation formulas give curves so smooth that parabolic interpolation between the points is simple and legitimate. Dr. Rhodes' method of graduating is not adapted to the elimination of seasonal fluctuations. It is also decidedly erratic in its results. Even with the short cuts Dr. Rhodes introduces, the calculation is laborious. We did not seriously consider using it in the interest rate and security price study.[2]

For a short series of data (such as the mortality figures used by Dr. Rhodes), which are not affected by seasonal fluctuations, a Whittaker-Henderson graduation is ideal.[3] From its very nature that method of smoothing gives perfect results if tested by the sum of the squares of the deviations

[1] Compare Rhodes, p. 8.

[2] The results obtained by Dr. Rhodes' method are slightly affected by the decision as to which end of the data shall be used as a beginning for operations. The summation formulas, and indeed all other formulas which are discussed in this book, give identical results whether they are worked forward or backward.

[3] Column VIII of Appendix V and Chart II give the results of applying such a graduation to Dr. Rhodes' test data. In this particular case $n = 2$. The significance of this statement will be understood when the immediately following text discussing the Whittaker-Henderson method of graduation has been read.

of the data from the graduation and the sum of the squares of the third differences of the graduation itself. When Henderson's method of computation is used, the labor involved in fitting to such a short series is not excessive. The sum of the squares of the 42 deviations of the data from the Whittaker-Henderson curve (with $n = 2$) is 359.17, which compares with 361.67 for the Rhodes 13-point curve, and 368.42 for the Rhodes 15-point curve. The sum of the squares of the third differences of the Whittaker-Henderson curve itself is 1.15, which compares with 1.61 for the Rhodes 15-point curve, and 5.58 for the Rhodes 13-point curve. The Whittaker-Henderson method of graduation is briefly explained in the immediately following chapter. The actual computation is outlined in Appendix VI.

CHAPTER VI

The Whittaker-Henderson Method of Graduation.

In smoothing data some departure from observed facts is, of course, necessary. In any particular case, the question naturally arises as to how far it is desirable to depart from the observed facts in order to obtain smoothness. If departure from the observed facts be measured by means of the sum of the squares of the deviations of the data from the graduated curve, and if roughness or lack of smoothness be measured by the sum of the squares of the third differences of the graduated curve, the problem is to decide at what stage it is undesirable to increase the sum of the squares of the deviations of the data from the graduated curve in order to decrease the sum of the squares of the third differences of the graduated curve itself.

Such a problem might be formulated as follows: How shall k times the sum of the squares of the deviations of the data from the graduated curve plus the sum of the squares of the third differences of that graduated curve be made a minimum? The larger the value chosen for k the more importance would be attached to closeness of fit as compared with smoothness. As the value of k approached infinity, the resulting graduated curve would approach the data points. As the value of k approached zero, the resulting graduated curve would approach a second degree parabola fitted to the observations by the method of least squares.[1]

In 1923, Professor E. T. Whittaker drew attention to the fact that if data be smoothed in such a manner that k times the sum of the squares of the deviations of the data from the smooth curve plus the sum of the squares of the third differences of the smooth curve itself be made a minimum, the smooth curve will be such that each of its ordinates will equal a corresponding ordinate of the data plus $\frac{1}{k}$ times a sixth difference of the smooth curve.[2] He developed a method for computing such a smooth curve.[3]

In 1924, Robert Henderson published an article in which he showed that if

[1] The form of the smooth curve in this limiting case is the result of making the sum of the squares of the *third* differences vanish and the sum of the squares of the deviations a minimum. If the sum of the squares of the *second* differences be made to vanish, the limiting case will give a straight line as the graduated curve.

[2] A complete set of sixth differences is obtained by extending the graduated curve by means of second-degree parabolas put through the first three and last three graduated points.

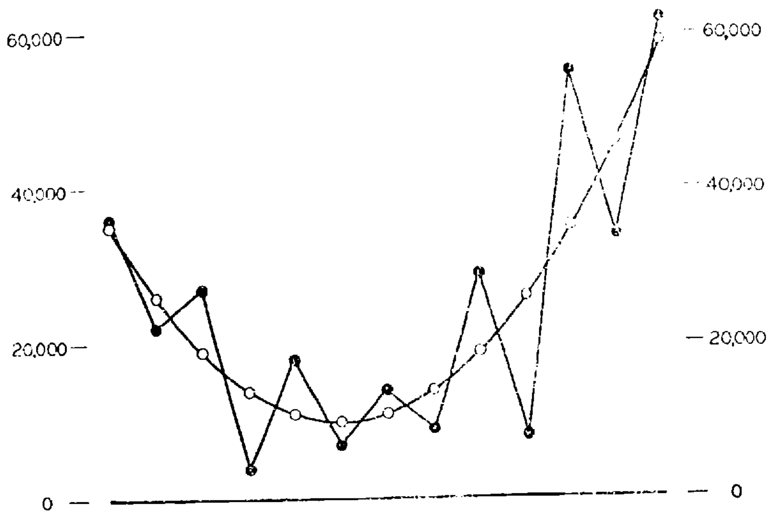
[3] See E. T. Whittaker and G. Robinson, *The Calculus of Observations*, pp. 303 to 315, inclusive. For an even earlier suggestion of the method, see Henderson and Sheppard, *Graduation of Mortality and Other Tables* (1919), page 6, eleven lines from bottom of page and onwards.

$$k = \frac{16\ (2n+3)^2}{n\ (n+1)^3\ (n+2)^3\ (n+3)},$$

the numerical calculation of the ordinates of the smooth curve can be arranged in a simple form for all values of k corresponding to integral values

CHART III

ILLUSTRATION OF
ROBERT HENDERSON'S
METHOD OF GRADUATION



of n.[1] While Henderson's calculation scheme is quite simple, the reader will probably understand it more easily if it is illustrated than he would if

[1] For the theory back of Henderson's method of fitting, see his article, *A New Method of Graduation*, Transactions of The Actuarial Society of America, May 1924, and Robert Henderson, *Some Points in the General Theory of Graduation*, Proceedings of the International Mathematical Congress—Toronto 1924—Volume II, pages 815-820, inclusive.

it were merely described. Appendix VI contains a numerical paradigm with complete instructions for computation. Chart III illustrates the fit obtained in Appendix VI.

Though Robert Henderson's method of computing the Whittaker-Henderson graduation is a highly elegant contribution to mathematical technique, and appears extremely simple, the actual computation is not without some serious pitfalls.

In the first place, the reader must remember that the three points from which the computation in the paradigm in Appendix VI began, were selected after the exact graduation was already known. They were selected in such a manner as to give immediately accurate results. In practice *the preliminary* operation (see Appendix VI) which is concerned with obtaining three points which shall closely approximate three points at one end of the true smooth curve, is often lengthy—particularly if the three points from which the preliminary operation commences are taken too close to the end or are at all badly chosen.[1]

Though each operation is extremely simple, being merely a multiplication by a small number or an addition or a subtraction, the operations have to be performed seriatim. In actual computation, this leads to errors being made with great ease. In any summation formula, each operation is com-

[1] See Note 1, page 152.

pleted before the next one is begun. For example, the first operation connected with the 43-term approximately fifth-degree parabolic formula is the taking of a 12-months moving total of *all* the data. The next step is the taking of an 8-months moving total of the first moving total—for *all* the data. Each type of operation is completed before the next is begun. Computers are much less likely to make mistakes if they do not have to change the nature of their operations momentarily.

Furthermore, while there are checks on the accuracy of the Henderson computation, such checks cannot be used until after the entire graduation has been calculated—unless we except the important check of repeating the preliminary operation backwards and forwards the entire length of the data until three identical, or practically identical, end figures have been arrived at. From the nature of the Henderson computation scheme errors are not easy to spot on a chart, as they tend to be distributed smoothly over a considerable range. This is not so with errors in some of the other graduations which have already been discussed. For example, errors in the computation of the parabolic summation formulas nearly always occur in the weight multiplications or in the final additions and subtractions of such results.[1] Now, such errors generally stand out as cusps on a chart. Moreover,

[1] The moving totals are automatically checked.

when discovered they can easily be corrected, whereas an error in the Henderson computation involves, at least theoretically, a re-computation of the whole *second half*.

If n equals 3, as in the paradigm in Appendix VI, the value of

$$\frac{4\,(2n+3)}{(n+1)\,(n+2)^2\,(n+3)}$$

is exactly .06. However, if n equals 2 or 4 or 5, for example, the value of the above function is no longer a simple decimal of few terms. The third figure of each column in the computation will then have to be obtained from a calculating machine.

It may well be that, for any particular set of observations, no integral value of n will give the exact graduation desired. If a fractional value of n be required, the computation becomes considerably lengthened.[1]

A Whittaker-Henderson graduation needs no extrapolation; it covers the entire range of the

[1] In 1925 Mr. Henderson published another article in which he described a method of applying the same graduation to a series of *weighted* observations. This second method he designated "Method B." Though "Method B" is primarily designed for use where the observations are weighted, it can be applied to unweighted observations by assuming a uniform weight of unity. However, when applied to unweighted observations, it has the disadvantage of being considerably more laborious in computation than Mr. Henderson's earlier method. On the other hand, it is more self-checking, and the difficulties connected with obtaining the three end points are taken care of in the method itself without involving any guessing as in the

data. This is a distinct element of mathematical elegance and sometimes an important practical consideration.[1] However, graduation of the ends of almost any series is necessarily extremely hypothetical unless facts outside the range covered by the graduation are used in obtaining the graduation. This is as true of the Whittaker-Henderson graduation as of any other type. If it were not so, we should have the philosopher's stone which turns all things to gold. A graduation could be applied to the course of stock or commodity prices, and successful speculation be based on the slope of the curve at its end point. Though mathematically inelegant, the most desirable procedure in a majority of the cases of graduation is to graduate not only the actual data, but extrapolated data which sometimes may be extremely crude estimates. However, we usually know *something* about the general nature of possible data outside of the range for which we have definite figures. It is highly undesirable not to make some use of such knowledge,

earlier method. Perhaps its outstanding superiority to the earlier method is that it is not restricted to any specific values of k. As in the earlier method, the graduation cannot be extended to include new data without re-calculation.

See Robert Henderson, *Further Remarks on Graduation*, Transactions of The Actuarial Society of America, 1925, pages 52 to 57 inclusive.

[1] Though there are simple mathematical methods of extrapolating any graduation in such a manner as to cover the entire range of the data, the labor involved in computation is, in most cases, large. See Chapter VIII.

indefinite though it may be.[1] While such a procedure may sometimes give a poorer fit, if tested by mathematical criteria based merely on the data covered by the graduation, it will tend to give a distinctly more rational fit. It will not allow any mere irregularity of the particular sample to determine the position of the graduation in the end regions.

Professor Whittaker stresses the fact that in obtaining the graduation all observations are used. The position of each datum point affects the position of every point on the smooth curve. While this is theoretically true, it is not of great practical importance. The alteration in the shape of the smooth curve at any point which would be caused by a change in a far distant datum point is negligible. The effect upon the smooth curve of changing a datum point is most pronounced in the immediate vicinity of the datum point. The results of such a change are less and less as the distance from the datum point increases. Moreover, it is highly desirable that such should be the case. It would be highly undesirable that a change in the position of a datum point should seriously affect the position of distant parts of the smooth

[1] One of the reasons why the Henderson graduations of the Call Money data, January 1886 to January 1894, look a little different from the other graduations at the ends is that, in computing the other graduations, data outside of the range, January 1886 to January 1894, were used. See Appendix VIII.

curve. For example, one of the great disadvantages of harmonic analysis is that the configuration of the data in one section may seriously affect the shape of the fitted curve in a far distant section. The Whittaker-Henderson graduation is free from this objection.

The Whittaker-Henderson graduation does not necessarily eliminate seasonal fluctuations. Only if n be taken sufficiently large will the percentage of seasonal fluctuation remaining be negligible. For example, if the seasonal fluctuation were a 12-months sine curve and if n equaled 5, over 95 per cent of the seasonal fluctuation would be eliminated. However, if n equaled 4, less than 89 per cent would be eliminated, and if n equaled 3, less than 69 per cent.

If k equal infinity (n equal zero) the graduation will coincide with the data no matter how irregular they may be. For all other values of k, the graduation "dampens" the data, no matter how smooth such data may be—with certain parabolic exceptions.[1]

If a Whittaker-Henderson graduation be applied to a regular and indefinitely extended sine series with complete period M the resulting smooth

[1] Strictly speaking a Whittaker-Henderson graduation in which n does not equal zero will not fall exactly on even a mathematical curve unless that curve be a second-degree parabola or a straight line.

curve will be a sine curve whose amplitude will be the amplitude of the original sine curve divided by

$$1 + \frac{4n\ (n+1)^3\ (n+2)^3\ (n+3)}{(2n+3)^2} \sin^6 \frac{\pi}{M}$$ [1]

The dampening effect on sine curves of various periods when n = 3, 4, or 5 may be seen from columns 25, 26 and 27 of the table in Appendix VII. It will be seen from that table that if we wish to eliminate the major portion of a 12-months sine seasonal we must make n as large as 5. Now if n = 5 all cycles whose period is less than 36 months are very inadequately fitted. A comparison of the figures for the 43-term graduation in column 24 of the same table will illustrate the point better than any lengthy discussion. The Whittaker-Henderson graduation is not particularly well adapted to saving the major portion of the amplitude of both short and long cycles, and at the same time eliminating seasonal fluctuations. This may be an important consideration in graduating such a series as monthly Call Money Rates.

Some of the purely practical disadvantages of the method have already been suggested. The computation is laborious—when account is taken of the difficulty of obtaining three accurate end points to begin *the first half* and the very considerable

[1] See Robert Henderson—*Discussion*—Transactions of The Actuarial Society of America, pages 306 and 307.

chance of error at all stages of the work. In using Method A, the computer does not have automatic checks as he proceeds with the work; if he use Method B, the advantage of checks is somewhat offset by the increased amount of computation involved. Upon the discovery of an error, the shortest procedure is usually to do the whole job over again. The graduation cannot be extended to include new data without re-calculating a large part of *the second half*—theoretically the whole of *the second half*. However, none of the above criticisms should be interpreted as in any way suggesting that the Henderson method is not a really superlative achievement in the theory and practice of graduation.

I wish to take this opportunity to thank Mr. Robert Henderson for his great kindness and courtesy in explaining his method of computation, answering numerous questions, and lending me illustrations of actual fittings.

CHAPTER VII

Desirable Characteristics of Formulas for Graduating Monthly Time Series.

A number of methods of mathematical smoothing have now been described and discussed. Most of these methods were found upon examination to be unsuited to describe such cyclical data as monthly Call Money Rates. However, two methods of graduating were found to give relatively excellent results when applied to such data. One of these methods is to use an approximately fifth-degree parabolic summation formula having simple computation weights. Three examples of such fifth-degree formulas were given—a 39-term formula whose computation weights are 10 and 15, a 43-term formula whose computation weights are 7 and 10, and a 33-term formula whose computation weights are 58 and 100.[1] The other method is to use the Whittaker-Henderson graduation outlined immediately above. For our purposes n would be taken in the Whittaker-Henderson grad-

[1] Third-degree or approximately third-degree parabolic formulas were suggested as desirable if the data series was very short or if the investigator wished to reduce the computation to a minimum. The 27-term formula described on page 28 was presented as the last word in ease of computation.

uation as equal to 5. Considerable preference was expressed for the first of these methods. We did not seriously consider using the Whittaker-Henderson method in spite of the mathematical elegancies inherent in that method. Two drawbacks to the use of the method were emphasized. It does not sufficiently eliminate seasonal fluctuations unless n be taken so large as to smooth the series more than we desired. The computation is not only laborious but its nature is such that mistakes are easily made.

In choosing a method of graduating the monthly series in the interest rate and security price study, two considerations were primary. First, the graduation must be good, that is, the graduated curve must not only be smooth but give a good fit to the data. Second, the computation must be easy, that is, it must not only be simple to understand but take little time to perform. It may be interesting to outline some characteristics which seemed most desirable in any formula to be used for graduating our particular time series.

1. *The graduation must be uninfluenced, or only negligibly influenced by distant observations:*

This requirement excluded any such procedure as harmonic analysis, unless such analysis be applied to successive portions of the curve in some sort of a moving manner. However, no such scheme was considered, as the resulting weight diagrams are not smooth and the computation is extremely laborious.

None of the graduations discussed in this book (with the possible exception of the Rhodes curve) are appreciably influenced by distant observations. Aside from Dr. Rhodes' graduation and the Whittaker-Henderson graduation, no point on any graduation in this book is in the slightest degree influenced by any observation further distant than 22 months.[1]

2. *The graduation must be easy to compute:*

We chose the summation type of computation with only the simplest multiplications. This type of computation is not only easy to perform but easy to understand. It is also extremely easy to check as the work proceeds. The entire elimination of multiplications is not of prime importance when the multiplications are so extremely simple as in the 39-term approximately fifth-degree parabolic formula described on page 71, the 43-term formula described on page 73, or the 33-term formula described on page 68.

[1] This is true even of graduations of data from which seasonal fluctuations have first been eliminated, if we do not consider the elimination of the seasonal fluctuation as a part of the graduation. See Appendix I.

3. *The weight diagram must be as smooth as possible:*

As the graduation was to be computed by means of a summation formula, it could be represented by a "weight diagram." This weight diagram should be as smooth as possible.[1]

4. *The graduation should eliminate 12-months seasonal fluctuations:*

As a summation formula was decided upon, the elimination of monthly seasonal fluctuations was obtained by having, as a possible first computing operation, the taking of a 12-months moving total of the data. All further operations are then to be performed on this 12-months moving total. All the formulas in this book which exactly eliminate 12-months seasonal fluctuations may be so calculated.

Of course a 12-months seasonal fluctuation may sometimes be of such a type that a graduation containing a 6-months or even a 3-months moving average will eliminate most of it. For example, the total dividend payments of American corpora-

[1] In spite of the above statement, the reader must not overemphasize the element of smoothness in the weight diagram. The weight diagram does not need to be superlatively smooth. He must remember that the increase in the smoothness of the graduation which results from using a superlatively smooth weight diagram instead of a merely ordinarily smooth one is negligible. Compare Note 1, page 56.

tions show a pronounced quarterly seasonal. A 3-months moving average will therefore remove a large percentage of the 12-months seasonal.

5. *If applied to successive points on a sine curve whose period is appreciably greater than the period of the seasonal fluctuation (in our case, 12 "points" or months), the graduation should fall as close as possible to the points on the sine curve:*

The exact fitting to sine curves of many periods cannot, of course, be rigidly fulfilled in practice. No formula, which would rigidly fulfill such a requirement, would, when applied to actual non-mathematical data, give a smooth curve. The element of compromise is introduced by the requirement of smoothness.

If any symmetrical set of weights be applied to an indefinitely extended sine curve, the resulting curve will itself be a sine curve, though not necessarily the same sine curve as the sine curve to which the weights have been applied. In certain limiting cases the resulting graduation may be a straight line. For example, the 43-term approximately fifth-degree parabolic summation formula,[1] if fitted to a sine curve whose period is 2, 3, 4, 5, 6, 8 or 12 months, or such a sine curve on which any straight line has been superposed, will necessarily give a straight line.

[1] See page 73.

For the graduation of most time series, it is much more important that the formula be capable of adequately describing various sine curves than capable of adequately describing any particular degree of parabola, even if such parabola be of as low a degree as the second. The insistence upon an absolute fit to any particular order of parabola is a statistical obsession. The 43-term approximately fifth-degree parabolic formula was designed primarily for fitting cyclical and not parabolic data, though it also gives an extremely close approximation to any parabola of a lower degree than the sixth.

Appendix VII, which is entitled "The Results of Applying Nineteen Different Graduation Formulas to Equidistant Points on Indefinitely Extended Sine Series," contains a table showing the percentage of the amplitudes of sine curves of various periods which are preserved by various graduation formulas. This table merits careful study, though the reader must remember that ability to fit sine curves of various periods is not the only characteristic which might be desired in a formula.

The first row in each column of the table in Appendix VII gives the goodness of fit to a 12-months sine curve. If the elimination of seasonal fluctuations is desired the entry in this column should be zero or close to zero. For this particular length of cycle, goodness of fit is *not* desired. For example, Spencer's 21-term formula which eliminates less than 45 per cent of a 12-months sine curve is distinctly not a formula to be used if seasonal elimination is desired.[1]

Smoothness of the resulting graduation is not considered in Appendix VII. Attention has already been drawn to smoothness of the weight diagram as an important factor leading to smoothness in the graduation. Another factor is the number of terms in the formula. Though the 43-term approximately fifth-degree parabolic formula [2] gives a distinctly closer fit to sine curves of different periods than does the 29-term non-parabolic formula.[3] it also gives a distinctly smoother graduation.

After the reader has studied Appendix VII showing the effects of applying various graduation formulas to sine curves, he should examine Appendix VIII. That Appendix gives the results of applying 14 different graduation formulas to the logarithms of 97 consecutive months of Call Money Rates on the New York Stock Exchange. The figures in Appendix VIII would well merit an adequate chart if it could be easily prepared. However, the various curves interweave so much that

[1] Unless the seasonal fluctuation be eliminated before graduation. See Appendix I.

[2] Column 24 of table in Appendix VII.

[3] Column 14 of table in Appendix VII.

it would be almost impossible to draw them all on one chart in such a manner as to show their respective characteristics unless the chart were made much larger than could be reproduced in this book. If the reader is sufficiently interested in the matter he may have a chart drawn showing the data and the various curves in colored inks. Such a chart, however, must be on an extremely large scale. One lying before me at the time I am writing is 20 inches wide and 39 inches long, and yet it is not on a scale large enough to carry all the 14 graduations without muddling the picture. A chart of about this size might be constructed on which the reader could examine any particular half dozen curves in which he is interested.

Chart IV shows the Call Money data, a 12-months moving average, and the 43-term approximately fifth-degree parabolic graduation.[1] Chart VI shows the data and two Whittaker-Henderson graduations, one with n = 3, the other with n = 5. Chart VII shows the data with a Spencer 21-term graduation and a Kenchington 27-term graduation. The reader should remember that the Whittaker-Henderson graduation with n = 3 eliminates less than 69 per cent of a 12-months sine curve seasonal and the Spencer 21-term graduation eliminates less than 45 per cent of such a seasonal. The

[1] The relation of Chart V to Chart IV is explained and discussed on page 25.

CHART IV

CALL MONEY RATES, TWELVE MONTHS MOVING AVERAGE AND 43-TERM CYCLICAL GRADUATION

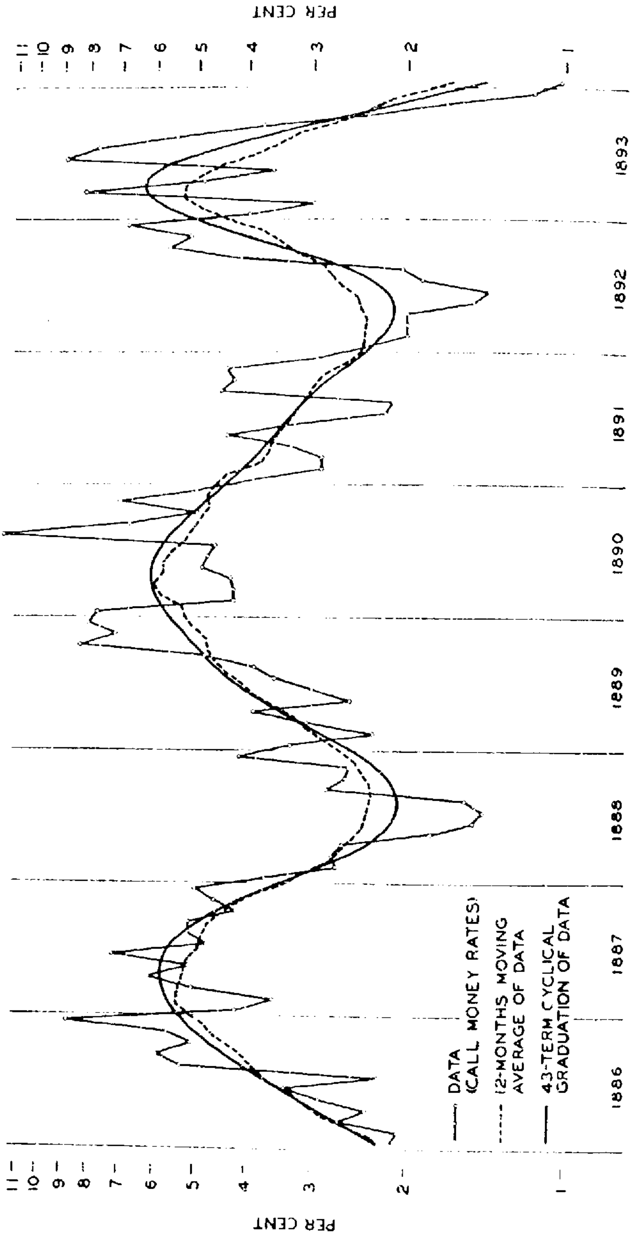


CHART V

A CRITERION OF GOODNESS OF FIT

TWELVE MONTHS MOVING AVERAGES OF THE THREE GRADUATIONS IN CHART IV

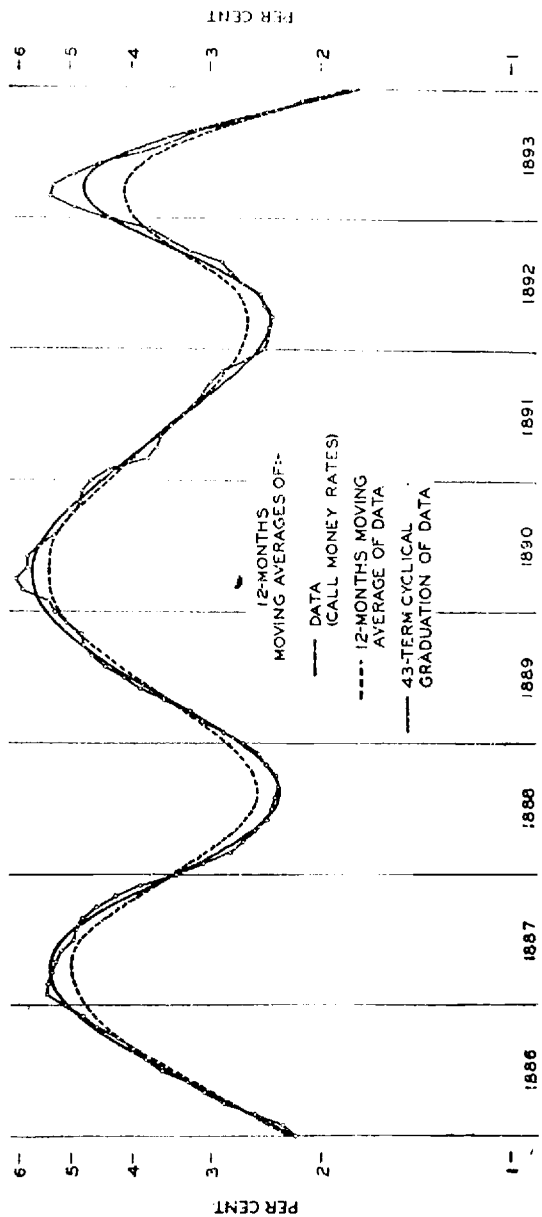


CHART VI

CALL MONEY RATES AND TWO WHITTAKER-HENDERSON GRADUATIONS

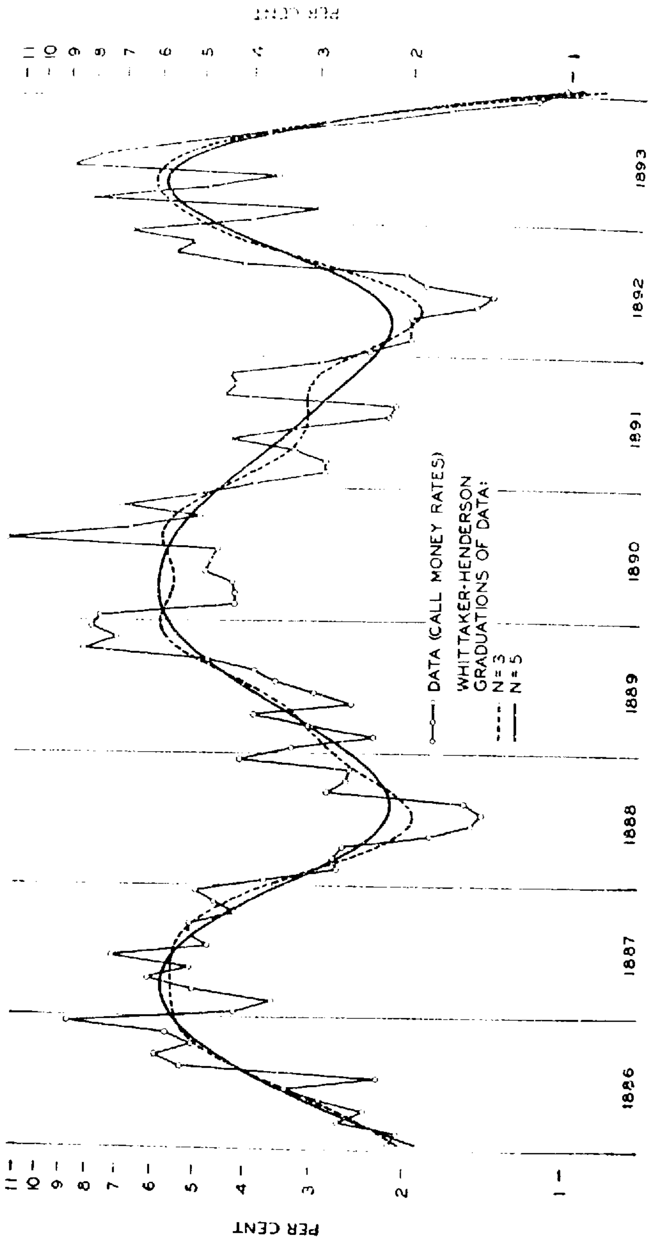
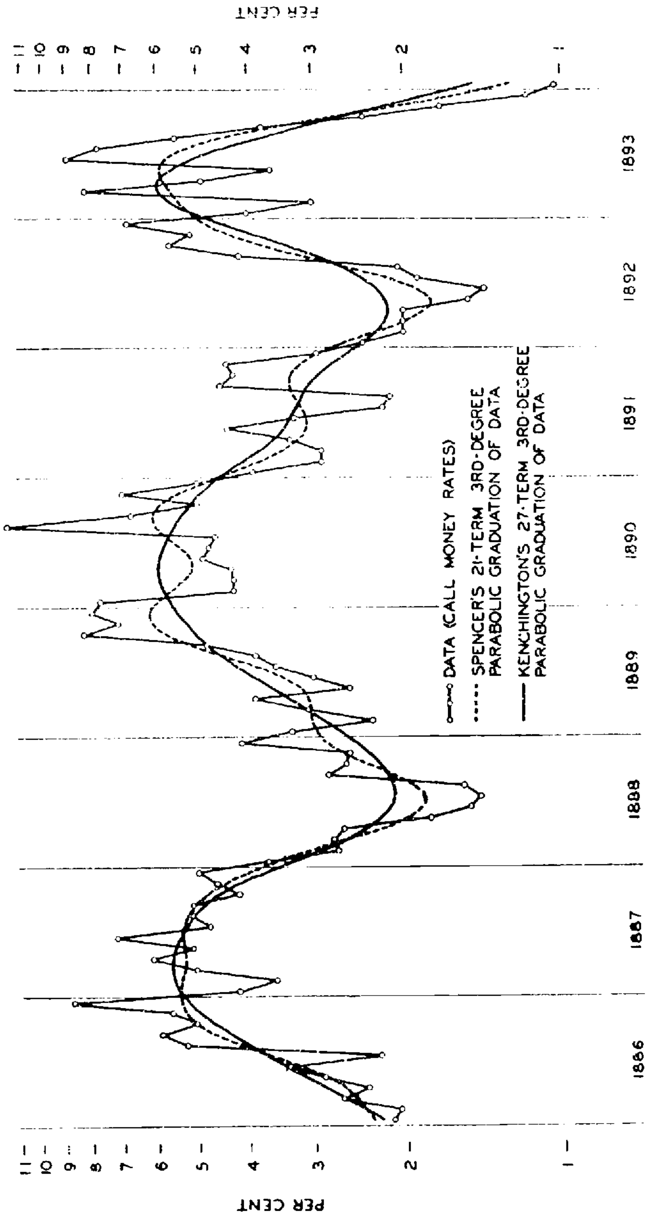


CHART VII

CALL MONEY RATES, KENCHINGTON'S 27-TERM GRADUATION AND SPENCER'S 21-TERM GRADUATION



Kenchington 27-term graduation eliminates over 90 per cent of a 12-months sine curve seasonal and the Whittaker-Henderson graduation, with $n = 5$, over 95 per cent of such a seasonal.

CHAPTER VIII

THE EXTENSION OF SUMMATION FORMULAS TO COVER THE RANGE OF THE DATA.

None of the graduations described in this book —except the Whittaker-Henderson graduations and the Rhodes curves—extends the entire length of the data. Extension of the graduation is always necessary. For example, if a 43-term graduation is to be made to cover the entire range of the data, 21 months at each end must be fitted by some other means. The case for extension of the graduation by means of hypothetical extrapolated data, outside the range of the actual data given, has already been briefly stated.[1]

If the investigator does not wish to use hypothetical data, simple mathematical methods of extrapolating the graduation without the use of further data are, of course, available. For example, a 43-term graduation may be made to cover the last 21 months by the following procedure: To the last 43 data points, fit a third-degree parabola by the method of least squares, in such a manner that it will have its middle, or 22nd, point coincident with the last point of the 43-term graduated curve,

[1] See pages 25, 26, 95, and 96, and Henderson, *Graduation of Mortality and Other Tables*, pages 37, 38 and 39.

and that its slope at this 22nd point will be the same as the slope of the 43-term curve at this last point of the 43-term curve. The slope of the 43-term curve at this last point is assumed to be the slope, at this point, of a second-degree parabola passing through the last 3 points of the 43-term curve.[1] It is better to fit a third-degree than a fourth-degree parabola. The chance of getting an absurd termination to the graduation is much greater with a fourth-degree parabola than with a third-degree.

[1] If the ordinates of the last 3 points of the 43-term curve be designated y_{-2}, y_{-1}, y_0, the slope at y_0 of the *second*-degree parabola $y = a + bx + cx^2$ drawn through these 3 points will be

$$\frac{y_{-2} - 4y_{-1} + 3y_0}{2}$$

If now a *third*-degree parabola ($y = A + Bx + Cx^2 + Dx^3$) be fitted to the last 43 data points by the method of least squares in such a manner that it goes through the last point of the 43-term curve, and has at this point the slope of the above fitted *second*-degree parabola at this point, the value of its 4 constants will be as follows:

$$A = y_0$$

$$B = \frac{y_{-2} - 4y_{-1} + 3y_0}{2}$$

$$C = \frac{\Sigma x^2 Y - y_0 \Sigma x^2}{\Sigma x^4}$$

$$D = \frac{\Sigma x^3 Y - B \Sigma x^4}{\Sigma x^6}$$

where y_{-2}, y_{-1}, y_0 are points on the 43-term curve and Y is an observational value (not a point on a smooth curve). The x's are, of course, points along the time axis. x_0 is the x of the 22nd point.

If the data contain a pronounced seasonal fluctuation, that fluctuation should be estimated by some good method (see Appendix I) and eliminated from the last 43 points of the data before the third-degree parabola is fitted to those data points. The extrapolation will then not be affected by mere seasonal elements in the data.

Such purely mathematical extrapolation is easier to defend than any method which involves judgment. The construction of hypothetical data, of course, always involves some judgment. However, the investigator must remember that it will not make a vast difference to the 43-term curve where the 21 extrapolated data points are placed, so long as they lie within a reasonable range.

The method is much less dangerous than any merely freehand extrapolation of the graduated curve itself. In a sense, the graduation is applied to data which at most are not quite half extrapolated. The procedure may be thought of as semi-mathematical. However, even a mere freehand extension of the graduation itself is not quite so bad as it sounds. The probable error of any fitted curve or graduation which covers the entire range of the data is bound to be large at the ends. A forecast of the future is inevitably implicit in the last few terms of the graduation. A friend of the writer once expressed himself on this matter. Said he, "I should certainly be willing to eat my hat if I could not forecast the future of Call Money Rates on the New York Stock Exchange more accurately than a third-degree parabola."

The forecast implicit in the mathematical extrapolation is bad as often as good. Indeed, the nature of a third-degree parabola is such that it seems as though the forecast were bad more often than good. If the last month for which we have observations be a panic month, such as November 1907, extension of the graduation by means of a third-degree parabola gives a curve rising more and more sharply, in other words, forecasting Call Money Rates at the end of 1908 as a large number of times the height they were in the high month of 1907. A third-degree parabola fitted to Railroad Stock Prices for the 43 months ending August 1909 would give a picture of an advancing market with no sign of hesitation. Of course, in this case, a hypothetical extension of the data might have been little better. The purely mathematical method may be used when the reader feels incapable of making any guess whatsoever. It can hardly then give worse results than introducing hypothetical data.

The argument in favor of purely mathematical methods of graduation which is derived from the fact that they are in general less laborious to apply than good judgment methods does not hold in the case of extending the graduation. The amount of computing necessary to cover the last points on the data by means of a fitted third-degree parabola is much greater than the amount necessary to extend the graduation by applying a summation formula to hypothetical data.

APPENDIX I

The Measurement of Probable Seasonal Fluctuations by Means of Operations on the Deviations of the Data from Graduated Curves. The Elimination of Seasonal Fluctuations before Graduation.

One object of the methods of graduation described in this book is the complete *elimination* of seasonal fluctuations. The problem of *measuring*—rather than *eliminating*—seasonal fluctuations has not been discussed. However, the problem of measurement must not be assumed necessarily divorced from that of elimination. Paradoxical as it may sound, the elimination of seasonal fluctuations is generally the best first step in the process of measuring them. For such measurement, operations on the deviations of the data from a smooth "seasonal eliminating" curve are theoretically more desirable than operations on the raw data. It is, of course, true that, as the number of years increases, extreme delicacy of smoothness or fit in the graduated curve becomes less and less practically important.

A few years ago the writer was approached by the statistical department of a government bureau and asked to propose a good but simple method of discovering any seasonal fluctuations which might exist in economic time series of moderate length. He replied that, as he did not know of any simple and yet really ideal method, he would suggest graduating the data roughly by means of a 2-months moving average of a 12-months moving average, taking the deviations of the data from this moving average (centered), and arriving at seasonal fluctuations from these deviations. Rough as is the method, it has been widely used and favorably noticed year after year. Moreover, though the method is extremely simple, in most cases the results are quite good. The averaging process used in obtaining a seasonal index for any one month generally makes unnecessary any great excellence in the graduation from which deviations are being measured.

A common method of obtaining seasonal fluctuations is to take the arithmetic average of each nominal month and then adjust for trend. This amounts to using a straight line as a base from which to measure deviations. If the number of years covered be large and if the assumption be made that the seasonal fluctuation is constant throughout the period, even such a crude and simple method often gives good results. Indeed, only if the number of years covered by the seasonal fluctuation is quite small, would it seem worth while to employ any particularly refined graduation—and only then if the erratic fluctuations were so small and the seasonal fluctuations so pronounced and regular as to make any deductions from such a short period legitimate.

Unless the number of years is very small, the results obtained from crude and from delicate graduations tend to be very similar. For example, the seasonal fluctuation of the 97 months of Call Money Rates from January 1886 to January 1894 (see Appendix VIII) is practically identical when computed from the *average* deviations of the data from a 43-term approximately fifth-degree parabolic graduation and when computed from the *average* deviations of the data from a 2-months moving average of a 12-months moving average.

The reader must remember that the arithmetic average of the *deviations* of the original data for successive Januaries from the successive January values given by the 43-term graduation equals the arithmetic average of the original data for successive Januaries minus the arithmetic average of the January values given by the 43-term graduation—similarly for other months than January and for the 2-months moving average of a 12-months moving average as well as for the 43-term graduation. Hence, only if the quasi-seasonal[1] obtained by

[1] The variations in the average values of the nominal months of either the 43-term graduation or the 2-months average of a 12-

averaging nominal months in the 43-term graduation differs appreciably from the quasi-seasonal obtained by averaging nominal months in the 2-months moving average of a 12-months moving average, will the seasonal obtained from *deviations* of the data from the 43-term graduation differ appreciably from the seasonal obtained from *deviations* of the data from the 2 of a 12-months moving average. Now an appreciable difference between the quasi-seasonal fluctuations of the 43-term graduation *itself* and quasi-seasonal fluctuations of the 2 of a 12-months moving average *itself* will occur only when the number of years covered is quite small—for data such as monthly Call Money Rates, say less than eight or nine years.

On the other hand, if the investigator wishes to make a careful study of *changing* seasonal fluctuations, he may well use some more delicate graduation than a 2 of a 12-months moving average, though it would seldom be worth while to use any formula involving much computation. The averaging process used in determining seasonal fluctuations will take care of any slight inadequacies in the smoothing formula. The 27-term formula re-

months moving average do not necessarily constitute any part of a true seasonal. Only by accident does any true seasonal remain in either of these graduations. There would be some variation in the average value of the nominal months of any curve whatsoever—except a straight line.

ferred to on page 28 would seem a highly desirable one to use. It is the last word in simplicity of calculation.[1]

Sometimes it is desirable to obtain an extremely close fit to all the factors in the data except seasonal fluctuations. A method of accomplishing this result is to eliminate seasonal fluctuations from the original data before final smoothing. This permits the use of non-seasonal-eliminating graduation formulas, which will follow the adjusted data extremely closely. For example, a Spencer 15-term formula may be applied to the data after the elimination of seasonal fluctuations and the resulting graduation will only accidentally contain any of the original seasonal fluctuations in spite of the closeness of its fit to all the minor movements of the adjusted data.

Charts VIII and IX show the results of applying a Spencer 15-term formula to the 97 months of Call Money data after adjustment for a constant seasonal fluctuation. On each chart is given a 43-term graduation for purposes of comparison. Chart VIII shows the two graduations and the data *after* adjustment for seasonal fluctuations. Chart IX

[1] Take a 16-months moving average of the data with the following simple weights: -1, 0, 0, 0, $+1$, $+1$, $+1$, $+1$, $+1$, $+1$, $+1$, $+1$, 0, 0, 0, -1. Take a 12-months moving total of this weighted 16-months moving total. Divide each of the final results by 72.

CHART VIII

CALL MONEY RATES ADJUSTED FOR SEASONAL FLUCTUATIONS,
43-TERM CYCLICAL GRADUATION
AND SPENCER'S 15-TERM GRADUATION OF THE ADJUSTED DATA

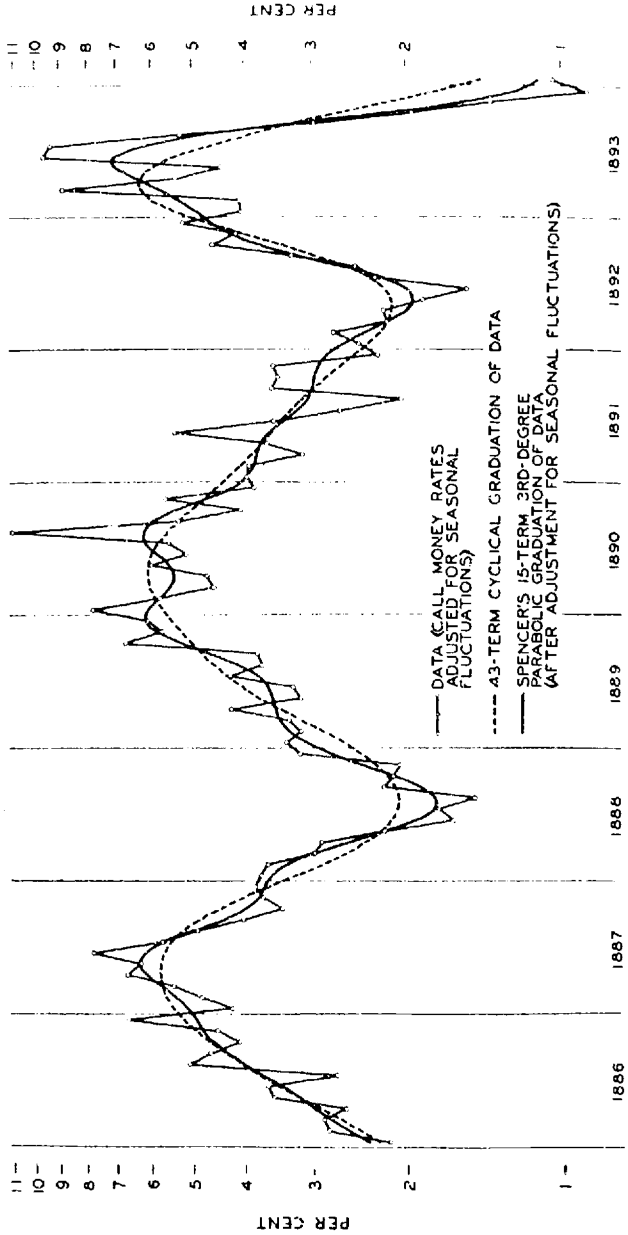
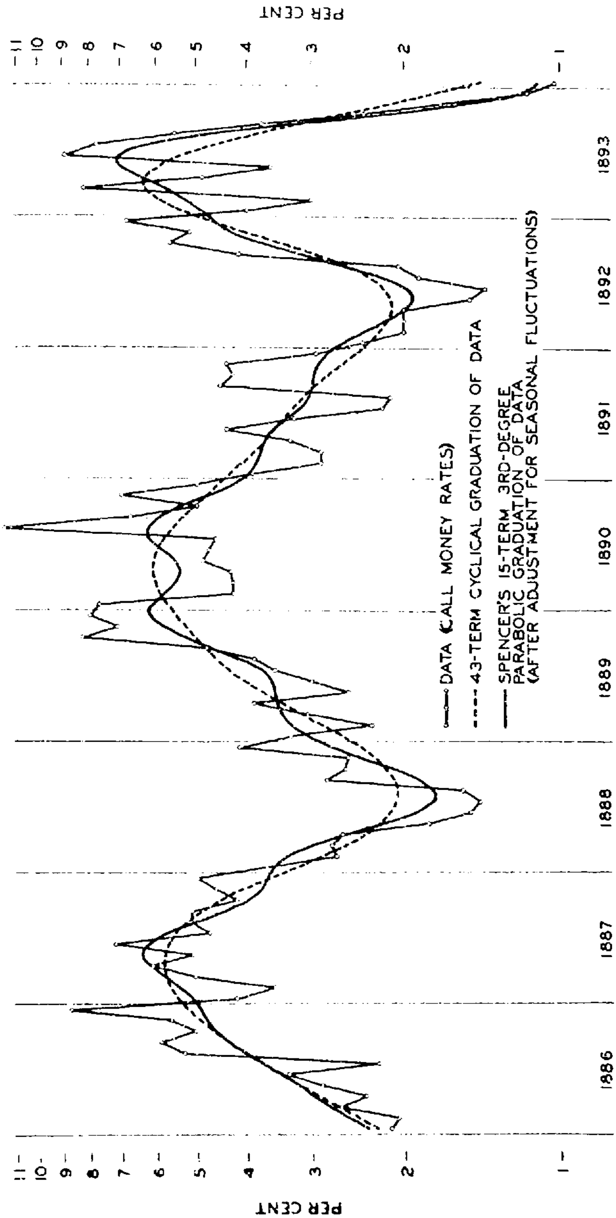


CHART IX

CALL MONEY RATES (NOT ADJUSTED FOR SEASONAL FLUCTUATIONS),
43-TERM CYCLICAL GRADUATION
AND SPENCER'S 15-TERM GRADUATION OF ADJUSTED DATA



shows the same two graduations and the data *before* adjustment for seasonal fluctuations.[1]

The *heavy* solid line in Chart X shows a Spencer 15-term graduation applied to the *adjusted* data. This is the same graduation as that shown in Charts VIII and IX. The broken line in Chart X shows a Spencer 15-term graduation applied to the *unadjusted* data. A moment's inspection will show that, with such pronouncedly seasonal data as Call Money Rates, the results of fitting such a graduation as Spencer's 15-term to the adjusted data are quite different from the results of fitting to the unadjusted data. The "close fit" of the Spencer 15-term graduation shown in Chart IX is a close fit to the "adjusted" data. Only if the adjustment is reasonable is the fit to the unadjusted data reasonable. The degree of reasonableness of the graduation of the unadjusted data depends upon the degree of reasonableness in the "seasonal" which has been eliminated.

If the seasonal fluctuations be uniform from year to year, the whole procedure is quite defensible. If the seasonal fluctuations actually eliminated be little more than some sort of an average of radically unlike movements, both the adjust-

[1] In both Chart VIII and Chart IX, the Spencer 15-term graduation is that obtained by graduating the *adjusted* data. Of course, no such statement need be made in connection with the 43-term graduation, as it gives the same results whether applied to adjusted or unadjusted data.

ment for seasonal fluctuations and the graduation of the adjusted data become difficult to defend. If the seasonal fluctuations seem to be, as in the case of the 97 months of Call Money Rates shown in Charts VIII, IX, and X, moderately though not pronouncedly regular, the graduation of the data. after adjustment for any "average" seasonal, is open to some slight criticism. It might be contended that it is a graduation of data which in sections has been adjusted for a seasonal fluctuation not existing in those sections.

Such a criticism cannot be made of the 43-term graduation. That graduation eliminates all seasonal fluctuation. It is unaffected by the reality or unreality of such seasonal fluctuation. It gives identically the same results when fitted to the unadjusted data as it does when fitted to the data adjusted for any seasonal fluctuation whatever—real or unreal. The adjustment of the data for an absolutely non-existent seasonal does not affect the results of applying the 43-term formula.

If seasonal fluctuations are to be eliminated before graduation, it is, of course, highly desirable that the investigator, before eliminating them, consider carefully whether they exist. For example, Call Money Rates since 1915 show a greatly reduced monthly seasonal fluctuation. To eliminate from such rates a seasonal derived from earlier rates would be quite illegitimate. A Spencer 15-

DESCRIPTION OF CHART X

Chart X is in two parts. The upper part shows Call Money Rates (unadjusted for seasonal fluctuations) and three Spencer 15-term graduations. The lower part of the chart shows a moving seasonal—on the same scale as the upper part of the chart.

The three Spencer 15-term graduations shown in the upper part of the chart are graduations of three different sets of data. The broken line is a graduation of the data as shown on the chart (unadjusted for seasonal fluctuations). The *heavy* solid line is a graduation of the data after adjustment for a constant seasonal. The *light* solid line is a graduation of the data after adjustment for the moving seasonal shown at the bottom of the chart.

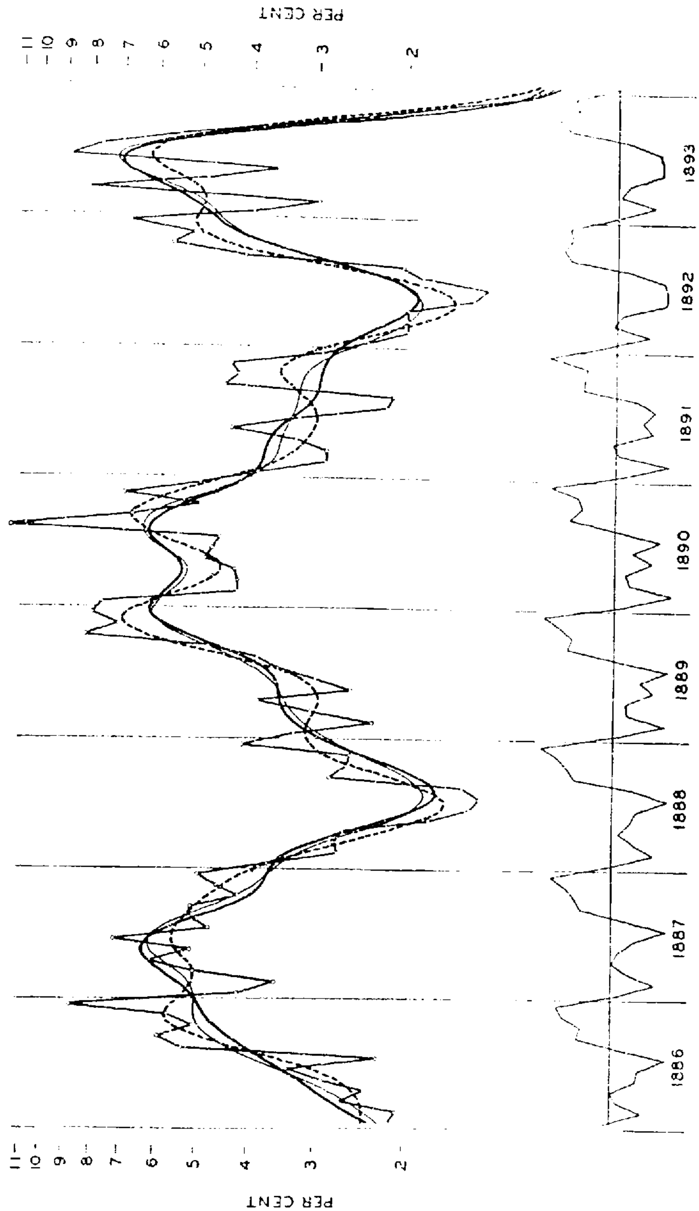
The constant seasonal was calculated by taking the arithmetic average of the deviations of each nominal month in the 97 from the 43-term cyclical graduation and adjusting for trend and to zero.

The construction of the moving seasonal may be illustrated as follows. The deviations of nine consecutive Januaries from the 43-term graduation were listed. The largest and the smallest deviations were then thrown out. The arithmetic average of the remaining seven deviations was taken. For example, the January 1886 seasonal was the arithmetic average of seven January deviations out of the nine Januaries from January 1882 to January 1890 inclusive. A similar procedure for each month gave a moving seasonal. This moving seasonal was corrected for sum by taking a 2-months moving average of a 12-months moving average of the seasonal and calling the deviations of the first moving seasonal from this graduation the final moving seasonal. As the trend is necessarily linear such a straight line formula was admissible.

In the calculations for both the constant seasonal and the moving seasonal, deviations were taken from the 43-term graduation because that graduation had already been computed. A much simpler formula, such as the 27-term formula described on page 28, would give perfectly satisfactory results.

CHART X

FOR DESCRIPTION OF THIS CHART SEE OPPOSITE PAGE



term formula applied to the monthly rates since 1915, after they had been adjusted for such a defunct seasonal, would give totally meaningless results. From 1857 to date, the changes in the characteristics of the seasonal fluctuations of Call Money Rates have been great. Generally the changes have been gradual though sometimes they have been sudden. When they have been sudden, the date of the change has usually corresponded with the date of some outstanding economic occurrence—such as the panic of 1873 or the early years of the Federal Reserve System. In such cases a particular seasonal fluctuation may be used up to a particular date, when another seasonal fluctuation will be substituted.

When the nature of the seasonal fluctuation appears to be changing gradually, the natural procedure would seem to be to calculate a *changing* seasonal. At the bottom of Chart X is a picture of such a changing seasonal. The solid *light* line in the upper part of the same chart gives the results of applying a Spencer 15-term graduation to the data after adjustment for this *changing* seasonal. The solid *heavy* line gives the results of applying a Spencer 15-term formula to the data after the elimination of a *constant* seasonal.[1] An examination of these two lines will show how much differ-

[1] For a description of the methods used in obtaining the constant seasonal and the changing seasonal, see page 130.

ence may appear in the graduation because of the elimination of different seasonal fluctuations—even in a period such as January 1886 to January 1894, when changes in the nature of the seasonal were not particularly violent.

If a less sensitive formula than Spencer's 15-term formula were used, the differences in the graduation resulting from differences in the seasonal fluctuations would, of course, be less. For example, Spencer's 21-term formula would be less affected by differences in the seasonal fluctuations than would Spencer's 15-term formula.[1]

[1] Spencer's 15-term formula is quite sensitive. Though it gives a graduation which is, of course, much smoother than that given by a simple 5-months moving average, it fits the data as closely as does the 5-months moving average.

If a formula with only a little greater smoothing power than the Spencer 15-term formula be desired, the following 17-term formula may be used: Take a 12-months moving total of the data with the following weights: -1, 0, 0, $+2$, $+2$, $+2$, $+2$, $+2$, $+2$, 0, 0, -1. (It will be noted that the plus weights constitute two times a simple 6-months moving total.) Take a 2-months moving total of a 5-months moving total of the results. As the total equals 100, division may be performed by merely moving the decimal point. The weight diagram is excellent. This is a very desirable formula, and the reader should not be disturbed by the fact that the graduation falls $\frac{3}{10}$ *outside* the parabola.

If, however, he desires a closer fit to a second degree parabola, he may use the following: Take a 14-months moving total with the following weights: -1, 0, 0, $+1$, $+2$, $+3$, $+4$, $+4$, $+3$, $+2$, $+1$, 0, 0, -1. (The plus weights constitute a 4-months moving total of a 5-months moving total.) Take a 2-months moving total of a 3-months moving total of the results. Divide by 108. The weight diagram is excellent. Falls $\frac{1}{6}$ *outside* the parabola $y = x^2$.

The sum of the squares of the third differences of the weights is, in each of the above formulas, much smaller than in Higham's

A crude, though extremely simple, procedure for eliminating seasonal fluctuations is sometimes used by financial writers and the editors of financial magazines and newspapers. The quotation for the present month is compared with the quotation for the same month in the preceding year. This comparison is made either by subtracting the quotation for the same month in the preceding year from the quotation for the present month or by dividing the quotation for the present month by the quotation for the same month in the preceding year.

The mathematical significance of any such operation may be described in a multitude of ways. Some ways are enlightening, others are not. There is a simple and enlightening way to describe the operation of *subtracting* the quotation for the same month last year from the quotation for the present month and using the resulting figure instead of the raw data. It amounts to taking a 12-months moving total of the data and using the first differences of this moving total instead of the raw data. The

17-term strictly third-degree parabolic formula—a 5 of a 5 of a 5 of $-1, +1, +1, +1, -1$ divided by 125.

If, on the other hand, a formula is desired which will follow the data even more closely than Spencer's 15-term formula, the following simple 13-term formula may be used: Take an 11-months moving total with the following weights: $-1, 0, +1, +2, +3, +4, +3, +2, +1, 0, -1$. (It will be noted that the plus weights constitute a 4-months moving total of a 4-months moving total.) Take a 3-months moving total of the results. Divide by 42. When applied to $y = x^2$, falls $\frac{1}{21}$ *outside* the parabola. The weight diagram is comparatively well-shaped.

procedure which consists of *dividing* the quotation for the present month by the quotation for the same month last year amounts to using the anti-logarithms of the first differences of a 12-months moving total of the logarithms of the data—instead of the raw data.

In either case the results are based upon month to month changes (first differences) of a crude graduation, namely, a 12-months moving average. Hence, even if an extremely good method of graduation were used instead of a 12-months moving average, the results would still be of the nature of a first derivative of the graduation. Moreover, as the 12-months moving average does not extend to the end of the data, its first differences do not tell whether, at the present time, the underlying curve of the data is high or low or whether it is rising or falling, but simply *whether it was rising or falling six months ago.*

Now, if the cycle were of unchanging period and shape such information would tell us something definite about the present. For example, if the data were a sine curve of 24 months period, we would know that, when the ratio of the present month to the same month last year began to increase, the sine curve had just passed through a low. The low of a 24-months sine curve occurs six months later than the point of inflection on the downward movement. However, if the data were a 48-months sine curve, we would know that when the ratio of the present month to the same month last year began to increase, the bottom of the sine curve was still six months ahead.

If the underlying curves are not sine curves, but less simple curves, the conclusions to be derived from the actions of such comparison of the present month with the same month last year must be considered most carefully. Though this type of operation on the data sometimes yields interesting results, such results must be carefully interpreted. The fact that the final figures are of the nature of a derivative curve must never be forgotten. The procedure seems likely to lead to misunderstandings when its results are intended for general public information.

APPENDIX II

A Computation Sheet Illustrating Graduation by Spencer's 15-Term Third-Degree Parabolic Formula.

Some readers of this book may be interested in examining a computation sheet for calculating a summation graduation. A computation sheet for graduating Rhodes' data (see Appendix V) by means of Spencer's 15-term formula (see page 55) appears below. The moving totals are calculated before applying the weights. This procedure makes the discovery of errors much easier. The moving totals are self-checking. Any mistake in the weight multiplications or their additions tends to stand out on the graduation like a sore thumb.

There are eleven columns in the paradigm given below. Columns 6, 8 and 9 may, of course, be eliminated, reducing the total number of columns to eight. This is, however, not desirable. The extra concentration needed in computation more than offsets the labor involved in copying out these columns as in the paradigm given below.

The columns in the computation sheet are:

1. Dr. Rhodes' data.
2. 5-year moving totals of column 1.
3. 4-year moving totals of column 2.

A Spencer 15-Term Graduation

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
137										
137										
131	669									
131	670									
133	661	2654								
138	654	2640	10512							
128	655	2619	10434							
124	649	2599	10350	31302	30798	41400	− 31536	− 30564	41400	129.38
132	641	2576	10266	31050	30564	41064	− 31302	− 30378	40998	128.12
127	631	2556	10188	30798	30378	40752	− 31050	− 30252	40626	126.96
130	635	2535	10126	30564	30252	40504	− 30798	− 30207	40315	125.98
118	628	2521	10084	30378	30207	40336	− 30564	− 30225	40132	125.41
128	627	2514	10069	30252	30225	40276	− 30378	− 30300	40075	125.23
125	624	2514	10075	30207	30300	40300	− 30252	− 30402	40153	125.48
126	635	2520	10100	30225	30402	40400	− 30207	− 30537	40283	125.88
127	634	2527	10134	30300	30537	40536	− 30225	− 30720	40428	126.34
129	634	2539	10179	30402	30720	40716	− 30300	− 30948	40590	126.84
127	636	2548	10240	30537	30948	40960	− 30402	− 31197	40846	127.64
125	644	2565	10316	30720	31197	41264	− 30537	− 31416	41228	128.84
128	651	2588	10399	30948	31416	41596	− 30720	− 31542	41698	130.31
		2615								

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
135	657	2631	10472	31197	31542	41888	− 30948	− 31521	42158	131.74
136	663	2638	10514	31416	31521	42056	− 31197	− 31371	42425	132.58
133	660	2630	10507	31542	31371	42028	− 31416	− 31119	42406	132.52
131	658	2608	10457	31521	31119	41828	− 31542	− 30816	42110	131.59
125	649	2581	10373	31371	30816	41492	− 31521	− 30504	41654	130.17
133	641	2554	10272	31119	30504	41088	− 31371	− 30201	41139	128.56
127	633	2529	10168	30816	30201	40672	− 31119	− 29886	40684	127.14
125	631	2504	10067	30504	29886	40268	− 30816	− 29550	40292	125.91
123	624	2480	9962	30201	29550	39848	− 30504	− 29160	39935	124.80
123	616	2449	9850	29886	29160	39400	− 30201	− 28680	39565	123.64
120	609	2417	9720	29550	28680	38880	− 29886	− 28119	39105	122.20
119	600	2374	9560	29160	28119	38240	− 29550	− 27462	38507	120.33
118	592	2320	9373	28680	27462	37492	− 29160	− 26742	37732	117.91
114	573	2262	9154	28119	26742	36616	− 28680	− 26001	36796	114.99
115	555	2198	8914	27462	26001	35656	− 28119	− 25269	35731	111.66
107	542	2134	8667							
101	528	2073	8423							
105	509	2018								
100	494									
96	487									
92										
94										

4. 4-year moving totals of column 3.
5. +3 times the values in column 4. The first figure 31302 equals the second figure of column 4 multiplied by 3—i.e., 10434 × 3. Etc.
6. +3 times the values in column 4. The first figure 30798 equals the fourth figure of column 4 multiplied by 3—i.e., 10266 × 3. Etc.
7. +4 times the values in column 4. The first figure 41400 equals the third figure of column 4 multiplied by 4—i.e., 10350 × 4. Etc.
8. −3 times the values in column 4. The first figure −31536 equals the first figure of column 4 multiplied by −3—i.e., 10512 × −3. Etc.
9. −3 times the values in column 4. The first figure −30564 equals the fifth figure of column 4 multiplied by −3—i.e., 10188 × −3. Etc.
10. Algebraic totals of columns 5, 6, 7, 8, 9.
11. Column 10 divided by 320.

APPENDIX III

Three Sets of Twenty-five Term Third-Degree Parabolic Graduation Weights.

For the method of obtaining the weights of column I, see page 54. The weights do not eliminate 12-months seasonal fluctuations. For the appearance of the weight diagram, see Figure 7, Chart I.

The weights of column II, if fitted to a second (or third) degree parabola, exactly fit the parabola. They eliminate 12-months seasonal fluctuations. The sum of the squares of the third differences of the weights is the minimum possible with the preceding restrictions. See page 58 and Figure 15, Chart I.

The weights of column III are rough approximations to those of column II. The resulting graduation falls $\frac{1}{6}$ (a negligible distance) *outside* the parabola $y = x^2$. It eliminates 12-months seasonal fluctuations. It is computed by taking $\frac{1}{144}$ of a 4-months moving total of a 12-months moving total of an 11-months moving total which has the following eleven simple weights: -1, 0, 0, $+1$, $+1$, $+1$, $+1$, $+1$, 0, 0, -1.

	I Henderson Ideal Third-Degree Parabolic	II Ideal Seasonal Eliminating Parabolic	III Summation Seasonal Eliminating Parabolic
	−.00334	−.00740	−.00695
	−.00890	−.01676	−.01389
	−.01369	−.02028	−.02083
	−.01456	−.01700	−.02083
	−.00922	−.00462	−.00695
	+.00321	+.01607	+.01389
	+.02217	+.04167	+.04167
	+.04580	+.06726	+.06944
	+.07138	+.08795	+.09028
	+.09572	+.10033	+.10417
	+.11576	+.10362	+.10417
	+.12892	+.10009	+.09722
	+.13350	+.09814	+.09722
	+.12892	+.10009	+.09722
	+.11576	+.10362	+.10417
	+.09572	+.10033	+.10417
	+.07138	+.08795	+.09028
	+.04580	+.06726	+.06944
	+.02217	+.04167	+.04167
	+.00321	+.01607	+.01389
	−.00922	−.00462	−.00695
	−.01456	−.01700	−.02083
	−.01369	−.02028	−.02083
	−.00890	−.01676	−.01389
	−.00334	−.00740	−.00695
Total	+1.00000	+1.00000	+1.00000

APPENDIX IV

The Weights Implied in Various Graduation Formulas.

In the text of this book a number of graduation formulas have been described and discussed as weighted moving averages. The present Appendix contains a table giving weights implied in fifteen of these graduation formulas. The "weight diagrams" for these fifteen graduations (with nine other graduations) are presented in Chart I (pages 77, 78, 79). The column numbers of Appendices IV, VII and VIII and the Figure numbers of Chart I are comparable. For example, Kenchington's 27-term formula is No. 12 in all three Appendices and in Chart I.

The weights given in this Appendix are those *implied* in the graduation formulas. For example, the weights implied in a 12-months simple moving average are 12 in number and all equal. As the total must equal unity, each weight equals 1/12. Chart I, Figure 1, represents such a system of weights. Each ordinate of Figure 1 is 1/12th of a unit high.

On pages 43 to 46, the possibility of describing various smoothing formulas as weighted moving averages was illustrated by discussing the "weights" implied in a 2-months moving average of a 12-months moving average (see Chart I, Figure 2), an 8-months moving average of a 12-months moving average (see Chart I, Figure 3), a 4-months moving average of a 5-months moving average of a 6-months moving average (see Chart I, Figure 4), a set of 13 weights such that, if applied to 13 consecutive and equally spaced observations, the result is the mid ordinate of a third-degree parabola fitted by the method of least squares (see Chart I, Figure 5).

The weight systems given in this Appendix are:

Col. 7. A Henderson Ideal 23-term third-degree parabolic graduation—see page 58, and Appendix VII.

Col. 8. A Henderson Ideal 29-term third-degree parabolic graduation—see Appendix VII.

Col. 9. A Henderson Ideal 33-term third-degree parabolic graduation—see Appendix VII.

Col. 11. Spencer's 21-term summation third-degree parabolic graduation—see Appendices VII and VIII.

Col. 12. Kenchington's 27-term summation third-degree parabolic graduation—see Appendices VII and VIII.

Col. 13. A 29-term summation approximately third-degree parabolic graduation (if fitted to parabola $y = x^2$, falls $\frac{7}{8}$ outside). See Appendices VII and VIII.

Col. 14. A 29-term summation non-parabolic graduation (if fitted to parabola $y = x^2$ falls $3\frac{1}{2}$ outside). See Appendices VII and VIII.

Col. 15. A 25-term "Ideal" 12-months seasonal eliminating third-degree parabolic graduation. See page 58 and Appendix VII.

Col. 18. A 35-term summation 5th-degree parabolic graduation. See Appendices VII and VIII.

Col. 19. A 41-term summation fifth-degree parabolic graduation. See Appendices VII and VIII.

Col. 20. A 43-term summation fifth-degree parabolic graduation. See Appendices VII and VIII.

Col. 21. A 45-term summation fifth-degree parabolic graduation. See page 66 and Appendices VII and VIII.

Col. 22. Another 45-term summation fifth-degree parabolic graduation. See page 65 and Appendices VII and VIII.

Col. 23. A 39-term summation approximately fifth-degree parabolic graduation. See Appendices VII and VIII.

Col. 24. A 43-term summation approximately fifth-degree parabolic graduation—this is the graduation used in the study of interest rates and security prices to eliminate 12-months seasonal and minor erratic fluctuations. See Appendices VII and VIII.

The first five weight systems of this Appendix (columns 7, 8, 9, 11, 12) do not eliminate 12-months seasonal fluctuations; the remaining weight systems do eliminate such 12-months seasonal fluctuations.

THE WEIGHTS IMPLIED IN VARIOUS GRADUATION FORMULAS

(The various graduation formulas have been given the same numbers in Appendices IV, VII and VIII and in Chart I.)

(7)	(8)	(9)	(11)	(12)
Henderson 25-term Ideal 3rd-degree Parabolic	Henderson 29 term Ideal 3rd-degree Parabolic	Henderson 33 term Ideal 3rd-degree Parabolic	Spencer's 21-term 3rd-degree Parabolic	Kenchington's 27-term 3rd-degree Parabolic
		−.00140		
		−.00418		
	−.00211	−.00754		
	−.00602	−.01031		−.00260
−.00334	−.01018	−.01137		−.00779
−.00890	−.01271	−.00981		−.01299
−.01369	−.01193	−.00509	−.00286	−.01558
−.01456	−.00678	+.00294	−.00857	−.01299
−.00922	+.00311	+.01400	−.01429	−.00260
+.00321	+.01733	+.02745	−.01429	+.01299
+.02217	+.03484	+.04240	−.00571	+.03377
+.04580	+.05413	+.05774	+.01714	+.05714
+.07138	+.07338	+.07234	+.05143	+.07792
+.09572	+.09075	+.08506	+.09429	+.09351
+.11576	+.10455	+.09493	+.13429	+.10649
+.12892	+.11341	+.10118	+.16286	+.11429
+.13350	**+.11646**	**+.10332**	**+.17142**	**+.11688**
+.12892	+.11341	+.10118	+.16286	+.11429
+.11576	+.10455	+.09493	+.13429	+.10649
+.09572	+.09075	+.08506	+.09429	+.09351
+.07138	+.07338	+.07234	+.05143	+.07792
+.04580	+.05413	+.05774	+.01714	+.05714
+.02217	+.03484	+.04240	−.00571	+.03377
+.00321	+.01733	+.02745	−.01429	+.01299
−.00922	+.00311	+.01400	−.01429	−.00260
−.01456	−.00678	+.00294	−.00857	−.01299
−.01369	−.01193	−.00509	−.00286	−.01558
−.00890	−.01271	−.00981		−.01299
−.00334	−.01018	−.01137		−.00779
	−.00602	−.01031		−.00260
	−.00211	−.00754		
		−.00418		
		−.00140		

The total of each column is unity.

(13) 29-term Approximately 3rd-degree Parabolic	(14) 29-term Non-parabolic	(15) 25-term Ideal Seasonal Eliminating 3rd-degree Parabolic	(18) 35-term 5th-degree Parabolic	(19) 41-term 5th-degree Parabolic
				+.00086
				+.00216
				+.00352
			+.00264	+.00367
			+.00600	+.00261
			+.00818	−.00050
$-.0027\dot{7}$	−.00232		+.00460	−.00528
$-.0083\dot{3}$	−.00694		−.00471	−.01131
$-.0138\dot{8}$	−.01389	−.00740	−.01789	−.01670
$-.0166\dot{6}$	−.01852	−.01676	−.02849	−.01955
$-.0138\dot{8}$	−.01852	−.02028	−.03007	−.01795
$-.0055\dot{5}$	−.01157	−.01700	−.02071	−.01064
$+.0083\dot{3}$	+.00231	−.00462	−.00306	+.00151
+.02500	+.02083	+.01607	+.01836	+.01827
$+.0416\dot{6}$	+.04167	+.04167	+.04167	+.03815
$+.0583\dot{3}$	+.06250	+.06726	+.06233	+.05922
+.07500	+.08102	+.08795	+.08039	+.07835
$+.0888\dot{8}$	+.09491	+.10033	+.09587	+.09448
+.10000	+.10417	+.10362	+.10880	+.10656
$+.1083\dot{3}$	+.10880	+.10009	+.11653	+.11420
$\mathbf{+.1111\dot{1}}$	**+.11110**	**+.09814**	**+.11912**	**+.11674**
$+.1083\dot{3}$	+.10880	+.10009	+.11653	+.11420
+.10000	+.10417	+.10362	+.10880	+.10656
$+.0888\dot{8}$	+.09491	+.10033	+.09587	+.09448
+.07500	+.08102	+.08795	+.08039	+.07835
$+.0583\dot{3}$	+.06250	+.06726	+.06233	+.05922
$+.0416\dot{6}$	+.04167	+.04167	+.04167	+.03815
+.02500	+.02083	+.01607	+.01836	+.01827
$+.0083\dot{3}$	+.00231	−.00462	−.00306	+.00151
$-.0055\dot{5}$	−.01157	−.01700	−.02071	−.01064
$-.0138\dot{8}$	−.01852	−.02028	−.03007	−.01795
$-.0166\dot{6}$	−.01852	−.01676	−.02849	−.01955
$-.0138\dot{8}$	−.01389	−.00740	−.01789	−.01670
$-.0083\dot{3}$	−.00694		−.00471	−.01131
$-.0027\dot{7}$	−.00232		+.00460	−.00528
			+.00818	−.00050
			+.00600	+.00261
			+.00264	+.00367
				+.00352
				+.00216
				+.00086

The total of each column is unity.

(20)	(21)	(22)	(23)	(24)
43-term 5th-degree Parabolic	45-term 5th-degree Parabolic	45-term 5th-degree Parabolic	39-term Approximately 5th-degree Parabolic	43-term Approximately 5th degree Parabolic
	+.00020	+.00033		
+.00056	+.00069	+.00127		+.00073
+.00143	+.00151	+.00272		+.00187
+.00233	+.00237	+.00384	+.00139	+.00312
+.00301	+.00300	+.00350	+.00347	+.00417
+.00321	+.00273	+.00138	+.00556	+.00469
+.00152	+.00108	−.00236	+.00556	+.00292
−.00177	−.00226	−.00668	+.00347	−.00083
−.00642	−.00688	−.01082	−.00208	−.00625
−.01179	−.01197	−.01393	−.01042	−.01271
−.01639	−.01612	−.01523	−.01875	−.01854
−.01821	−.01771	−.01398	−.02431	−.02135
−.01603	−.01542	−.00961	−.02431	−.01979
−.00923	−.00835	−.00237	−.01806	−.01323
+.00292	+.00364	+.00796	−.00347	−.00063
+.01947	+.01986	+.02143	+.01597	+.01698
+.03865	+.03867	+.03816	+.03819	+.03750
+.05832	+.05837	+.05668	+.06042	+.05854
+.07746	+.07711	+.07501	+.08125	+.07917
+.09378	+.09324	+.09112	+.09792	+.09667
+.10579	+.10543	+.10343	+.10972	+.10937
+.11333	+.11302	+.11125	+.11806	+.11739
+.11612	**+.11558**	**+.11380**	**+.12084**	**+.12042**
+.11333	+.11302	+.11125	+.11806	+.11739
+.10579	+.10543	+.10343	+.10972	+.10937
+.09378	+.09324	+.09112	+.09792	+.09667
+.07746	+.07711	+.07501	+.08125	+.07917
+.05832	+.05837	+.05668	+.06042	+.05854
+.03865	+.03867	+.03816	+.03819	+.03750
+.01947	+.01986	+.02143	+.01597	+.01698
+.00292	+.00364	+.00796	−.00347	−.00063
−.00923	−.00835	−.00237	−.01806	−.01323
−.01603	−.01542	−.00961	−.02431	−.01979
−.01821	−.01771	−.01398	−.02431	−.02135
−.01639	−.01612	−.01523	−.01875	−.01854
−.01179	−.01197	−.01393	−.01042	−.01271
−.00642	−.00688	−.01082	−.00208	−.00625
−.00177	−.00226	−.00668	+.00347	−.00083
+.00152	+.00108	−.00236	+.00556	+.00292
+.00321	+.00273	+.00138	+.00556	+.00469
+.00301	+.00300	+.00350	+.00347	+.00417
+.00233	+.00237	+.00384	+.00139	+.00312
+.00143	+.00151	+.00272		+.00187
+.00056	+.00069	+.00127		+.00073
	+.00020	+.00033		

The total of each column is unity.

APPENDIX V

Seven Graduations of Dr. E. C. Rhodes' Test Mortality Data.

The data are rates of infantile mortality from causes other than diarrheal diseases, for the 42 years from 1870 to 1911, inclusive. See page 82, note 1.

GRADUATIONS OF DR. RHODES' TEST DATA

I	II	III	IV	V	VI	VII	VIII
Rhodes' Original Data	Rhodes' 15-point Curve	Rhodes' 13-point Curve	Rhodes' 11-point Curve	Rhodes' 9-point Curve	Shep-pard's Gradua-tion	Spencer's 15-term Formula	Whittaker-Henderson n = 2
137	137.2	137.7	137.7	138.5	136.5	137.11	136.72
137	135.0	134.6	134.3	132.7	135.4	135.10	135.27
131	133.6	133.1	133.1	132.7	134.3	133.77	134.07
131	132.8	132.6	132.9	134.1	133.2	132.84	133.11
133	132.1	132.4	132.8	134.2	132.2	132.07	132.27
138	131.5	132.0	132.1	132.2	131.2	131.29	131.37
128	130.7	131.2	130.7	128.7	130.2	130.41	130.31
124	129.7	130.0	129.2	125.6	129.2	129.38	129.19
132	128.5	128.4	127.8	125.3	128.4	128.12	128.11
127	127.2	126.7	126.7	127.1	127.1	126.96	127.08
130	126.1	125.5	126.0	128.1	125.7	125.98	126.17
118	125.4	124.7	125.2	127.7	125.2	125.41	125.50
128	125.2	124.3	124.6	127.0	126.1	125.23	125.21
125	125.3	124.5	125.0	126.3	126.0	125.48	125.27
126	125.6	125.4	126.2	125.5	125.0	125.88	125.62
127	126.1	126.5	127.4	124.6	125.6	126.34	126.20
129	126.8	127.3	127.7	124.5	126.7	126.84	126.98
127	127.8	128.1	127.6	125.3	128.8	127.64	127.96
125	129.0	129.2	128.0	127.0	129.7	128.84	129.15
128	130.4	130.8	129.3	130.6	130.3	130.31	130.45
135	131.8	132.3	131.4	134.8	131.2	131.74	131.57
136	132.7	133.1	133.0	137.2	131.7	132.58	132.17
133	133.1	133.0	133.2	135.4	131.9	132.52	132.08
131	132.6	131.9	132.1	130.9	131.7	131.59	131.37
125	131.3	130.1	130.4	126.9	130.5	130.17	130.25
133	129.5	128.0	128.9	125.0	129.1	128.56	128.94
127	127.6	126.3	127.4	124.7	127.2	127.14	127.55
125	125.9	125.1	125.8	124.9	126.0	125.91	126.19
123	124.4	124.1	124.1	125.4	124.7	124.80	124.89
123	122.9	123.1	122.7	126.3	123.8	123.64	123.57
126	121.3	121.9	121.5	126.0	122.6	122.20	122.04
119	119.5	120.7	120.2	122.9	119.5	120.33	120.12
118	117.4	118.9	118.4	117.8	117.5	117.91	117.75
114	114.9	116.1	115.8	112.7	115.0	114.99	114.95
115	112.1	112.7	112.6	108.4	112.1	111.66	111.79
107	109.1	109.1	109.1	105.1	108.3	108.09	108.40
101	105.9	105.4	105.5	103.4	105.5	104.55	105.00
105	102.7	101.8	101.9	103.2	102.6	101.21	101.78
100	99.6	98.6	98.6	102.7	99.8	98.24	98.83
96	96.7	96.1	95.7	100.7	97.0	95.82	96.30
92	94.2	94.7	93.8	96.7	94.2	94.11	94.36
94	92.2	94.9	93.0	90.7	91.4	93.30	93.15

APPENDIX VI

Paradigm for Graduating by Robert Henderson's Method.

In the immediately following paradigm n is taken as equal to 3. For general discussion of the method see Chapter VI.

Let there be 13 observations equally spaced on the x axis. Let the ordinates of the observations be 36009, 22009, 27018, 4027, 18045, 7054, 14054, 9045, 29036, 8027, 55036, 34036, and 62054. It is desired to graduate these observations in such a manner that $\frac{9}{1000}$[1] times the sum of the squares of the deviations of the data from the graduated curve plus the sum of the squares of the third differences of the graduated curve shall be a minimum. The actual computation is in three steps. Mr. Henderson calls the three steps, *the preliminary*, *the first half*, and *the second half*. The paradigm[2] on pages 155 and 156 should be followed when reading the instructions below.

[1] The reader will note from page 91 that $k = \frac{9}{1000}$ when $n = 3$.

[2] For the paradigm the data above were so chosen that all ordinates of the graduated curve would be integers. In actual computation, the calculations would be carried to as many decimals as were desired in the graduated curve.

The Preliminary:

1. *Guess* three points on the graduated curve some distance from the beginning. The three points in the paradigm were chosen in such a manner as to give immediate results in the graduation.[1] Their ordinates are 26000, 18907, and 14213. These values are supposed to be guesses at the 10th, 9th and 8th ordinates of the graduated curve.

2. Find the first and second differences of the above three ordinates. Use —4694 as the first difference (*i.e.*, 14213 — 18907). The second difference is +2399.
3. Calculate the figures in the paradigm by filling each column before beginning on the next. Beginning with the first figure of the first column, we notice:

   (a) +2399 is the second difference mentioned above.
   (b) —4694 is the first difference mentioned above.
   (c) +14213 is the guess at the 8th ordinate of the graduated curve. (See paragraph 1.)
   (d) —18776 = —4694 × 4. (See note 1, page 153.)
   (e) +23990 = +2399 × 10. (See note 1, page 153.)
   (f) +19427 = +14213 — 18776 + 23990
   (g) The first figure (+4027) of the second column is the 4th *datum* ordinate.
   (h) —15400 = +4027 — 19427

---

[1] The preliminary portion of the Henderson method of computation is used simply to obtain suitable estimated figures with which to begin *the first half*.

If the three preliminary points are badly chosen, the preliminary operation will be lengthened. In our illustration, it might have to be extended through *the first half* or even through another return operation. For final results accurate to any particular number of decimals, it would be necessary to work backwards and forwards until at two separate stages three values which were identical to the required number of decimals were obtained at one of the ends.

It is highly desirable to choose 3 preliminary values a considerable distance along the curve (say 20 units when n = 3 and more when n is larger) and to choose them as well as possible. Mr. Henderson advocates fitting a second-degree parabola to eleven consecutive points some distance along the curve, and using the 4th, 5th and 6th terms of the parabola as the three preliminary points.

(i) $-924 = -15400 \times .06$[1]
(j) $+1475 = -924 + 2399$
(k) $-3219 = +1475 - 4694$
(l) $+10994 = -3219 + 14213$
(m) $-12876 = -3219 \times 4$
et cetera

4. The twelve figures at the end of *the preliminary* are calculated as follows: The first three figures are the 6th figures of the last three columns. The other 9 figures are obtained from these 3 figures by assuming the 3 figures to lie on a second-degree parabola and extrapolating the parabola by means of first and second differences.

The First Half:

1. The first three figures of column one are obtained from the last three of the twelve figures just discussed in *the preliminary*.
   (a) +2009 is their second difference.
   (b) —19045 is a first difference. As we are now going backwards, it equals +91099 — 110144.
   (c) +91099 is the 10th figure of the last 12 of *the preliminary*.
2. The computation now runs as it did in *the preliminary* except that the first figure of the second column (+36009) is the first datum ordinate instead of the fourth as in *the preliminary*.
3. The last three figures (+35036, +46036, +59054) are extrapolated from +14144, +19090, +26054, by means of first and second differences.

The Second Half:

1. The reader will be able to understand *the second half* quite easily if he will notice that the figures at the tops of the

---

[1] If $n = 3$

$$n + 1 = 4$$

$$\frac{(n+1)\ (n+2)}{2} = 10$$

$$\frac{4\ (2n+3)}{(n+1)\ (n+2)^2\ (n+3)} = .06$$

For the theory back of all this, see the Henderson articles referred to in note 1, page 91.

columns are not data ordinates (as in *the first half*) but are the 6th figures of columns in *the first half* taken in reverse order.

2. The first three figures of the first column of *the second half* are obtained from the last three (extrapolated) figures of *the first half* in the same manner that the first three figures of *the first half* were obtained from *the preliminary*.
3. The 13 ordinates of the graduated curve are the sixth figures of the columns of *the second half* (taken in reverse order) to which are added the last three figures of *the first half*. They are +35009, +26009, +19018, +14027, +11045, +10054, +11054, +14045, +19036, +26027, +35036, +46036, +59054. If the reader will calculate the sixth differences of these figures, he will discover that each of them equals minus $\frac{9}{1000}$ times the corresponding deviation of a datum ordinate from the graduated curve, and hence the sum of the squares of the third differences of the graduated curve plus $\frac{9}{1000}$ times the sum of the squares of the deviations of the data from the graduated curve is a minimum.

The Paradigm

Paradigm for graduating data in such a manner that $\frac{9}{1000}$ times the sum of the squares of the deviations of the data from the graduation plus the sum of the squares of the third differences of the graduation shall be a minimum.[1]

[1] In the following paradigm data ordinates are italicized and ordinates of the final graduation are set in heavy full face type.

Preliminary

	+ 4027	*+27018*	*+22009*	*+36009*
	−15400	+14150	− 7750	+ 2500
	− 924	+ 849	− 465	+ 150
+ 2399	+ 1475	+ 2324	+ 1859	+ 2009
− 4694	− 3219	− 895	+ 964	+ 2973
+14213	+10994	+10099	+11063	+14036
−18776	−12876	− 3580	+ 3856	
+23990	+14750	+23240	+18590	
+19427	+12868	+29759	+33509	

+ 10099
+ 11063
+ 14036
+ 19018
+ 26009
+ 35009
+ 46018
+ 59036
+ 74063
+ 91099
+110144
+131198

First Half

	+36009	*+22009*	*+27018*	*+ 4027*	*+18045*
	+ 1000	− 4900	+11150	−15550	+13000
	+ 60	− 294	+ 669	− 933	+ 780
+ 2009	+ 2069	+ 1775	+ 2444	+ 1511	+ 2291
−19045	−16976	−15201	−12757	−11246	− 8955
+91099	+74123	+58922	+46165	+34919	+25964
−76180	−67904	−60804	−51028	−44984	−35820
+20090	+20690	+17750	+24440	+15110	+22910
+35009	+26909	+15868	+19577	+ 5045	+13054

+ 7054	*+14054*	*+ 9045*	*+29036*	*+ 8027*	*+55036*
− 6000	+ 3900	− 7250	+13300	−25200	+32450
− 360	+ 234	− 435	+ 798	− 1512	+ 1947
+ 1931	+ 2165	+ 1730	+ 2528	+ 1016	+ 2963
− 7024	− 4859	− 3129	− 601	+ 415	+ 3378
+18940	+14081	+10952	+10351	+10766	+14144
−28096	−19436	−12516	− 2404	+ 1660	+13512
+19310	+21650	+17300	+25280	+10160	+29630
+10154	+16295	+15736	+33227	+22586	+57286

+34036	*+62054*			
−23250	+ 7500			
− 1395	+ 450			
+ 1568	+ 2018			
+ 4946	+ 6964	+ 8982	+11000	+13018
+19090	+26054	**+35036**	**+46036**	**+59054**
+19784				
+15680				
+54554				

Second Half

	+10766	+10351	+10952	+14081	+18940
	− 450	+ 450	− 300	0	− 150
	− 27	+ 27	− 18	0	− 9
+ 2018	+ 1991	+ 2018	+ 2000	+ 2000	+ 1991
−11000	− 9009	− 6991	− 4991	− 2991	− 1000
+35036	**+26027**	**+19036**	**+14045**	**+11054**	**+10054**
−44000	−36036	−27964	−19964	−11964	− 4000
+20180	+19910	+20180	+20000	+20000	+19910
+11216	+ 9901	+11252	+14081	+19090	+25964

+25964	+34919	+46165	+58922	+74123
0	0	+ 300	− 150	+ 150
0	0	+ 18	− 9	+ 9
+ 1991	+ 1991	+ 2009	+ 2000	+ 2009
+ 991	+ 2982	+ 4991	+ 6991	+ 9000
+11045	**+14027**	**+19018**	**+26009**	**+35009**
+ 3964	+11928	+19964	+27964	
+19910	+19910	+20090	+20000	
+34919	+45865	+59072	+73973	

Summary

Data	*Graduation*
36009	35009
22009	26009
27018	19018
4027	14027
18045	11045
7054	10054
14054	11054
9045	14045
29036	19036
8027	26027
55036	35036
34036	46036
62054	59054

APPENDIX VII

The Results of Applying Nineteen Different Graduation Formulas to Equidistant Points on Indefinitely Extended Sine Series.

The entries in the table show the percentages of the amplitudes of the various sine curves which are preserved by each formula.

(The various graduation formulas have the same numbers in Appendices IV, VII and VIII and in Chart I.)

(0)	(2)	(7)	(8)	(9)	(11)
Sine Curve Period (Points)	2 of a 12 Simple Moving Average	Henderson 25-term Ideal 3rd-degree Parabolic	Henderson 29-term Ideal 3rd-degree Parabolic	Henderson 33-term Ideal 3rd-degree Parabolic	Spencer's 21-term 3rd-degree Parabolic
12	0	24.54	8.07	− 1.40	55.22
15	23.04	53.53	36.12	20.21	76.09
18	40.93	71.98	58.70	44.31	86.65
20	50.04	79.75	69.22	57.05	90.66
24	63.30	88.93	82.47	74.42	95.10
30	75.41	94.97	91.77	87.54	97.29
36	82.50	97.44	95.73	93.41	98.94
40	85.67	98.28	97.11	95.47	99.29
48	89.89	99.14	98.54	97.70	99.63
60	93.48	99.64	99.38	99.03	99.87
120	98.33	99.98	99.96	99.96	99.95

Notes:

Col. (0) *Sine curve periods.* The first entry in column 2 (zero) means that if the formula of column 2 be applied to equidistant monthly points on an indefinitely extended sine curve whose period is 12 months, such sine curve is entirely eliminated. The formula will give a horizontal straight line. The first entry in column (7) means that if the formula of column (7) be applied to such a 12-months sine curve, whose amplitude is 100 (vertical distance between minimum values and maximum values), the curve resulting from the application of the formula will be a 12-months sine curve whose amplitude will be 24.54.

Col. (2) See page 43.
Col. (7) See page 59.
Col. (8) See Appendix IV.
Col. (9) See page 57 and Appendix IV.
The first entry in this column (*minus* 1.40) may disturb the reader. It signifies that if this particular formula be applied to a 12-months sine curve the resulting smooth curve will be a sine curve whose amplitude will be 1.40 per cent of the amplitude of the original sine curve but which will have maxima where the original curve had minima and vice versa. It will slightly *overcorrect* for a 12-months sine seasonal.
Col. (11) See pages 51, 52, 53.

	(12)	(13)	(14)	(15)	(18)
SINE CURVE PERIOD (POINTS)	KENSINGTON'S 27-TERM 3RD-DEGREE PARABOLIC	29-TERM APPROXIMATELY 3RD-DEGREE PARABOLIC	29-TERM NON-PARABOLIC	25-TERM IDEAL SEASONAL ELIMINATING 3RD-DEGREE PARABOLIC	35-TERM 5TH-DEGREE PARABOLIC
12	9.88	0	0	0	0
15	40.86	29.27	33.48	35.29	45.18
18	63.07	54.56	60.05	60.02	72.81
20	72.90	66.52	72.06	70.81	82.90
24	84.88	81.60	86.56	83.83	92.89
30	93.03	92.03	95.86	92.59	97.78
36	96.42	96.32	99.24	96.21	99.20
40	97.58	97.13	100.19	97.44	99.55
48	98.78	99.12	100.92	98.71	99.82
60	99.50	99.86	101.05	99.47	99.97
120	99.94	100.09	100.40	99.95	99.98

Col. (12) See pages 29, 58.
Col. (13) $\frac{1}{360}$ of a 2-months moving total of a 12-months moving total of the results of subtracting a 17-months moving total from a 4-months moving total of an 8-months moving total. This formula is not *rigidly* parabolic. It falls $\frac{7}{6}$ outside the parabola $y = x^2$. See pages 59 and 60.
Col. (14) $\frac{1}{432}$ of a 3-months moving total of a 12-months moving total of the results of subtracting a 16-months moving total from a 4-months moving total of a 7-months moving total. Falls $3\frac{1}{2}$ outside parabola $y = x^2$. See pages 27, 60.
Col. (15) See page 58.
Col. (18) See pages 67, 68.

	(19)	(20)	(21)	(22)	(23)
Sine Curve Period (Points)	41-term 5th-degree Parabolic	43-term 5th-degree Parabolic	45-term 5th-degree Parabolic	45-term 5th-degree Parabolic	39-term Approximately 5th-degree Parabolic
12	0	0	0	0	0
15	36.41	34.16	33.13	26.89	43.03
18	65.34	63.22	62.21	56.55	73.25
20	77.27	75.60	74.80	70.41	84.28
24	90.01	89.10	88.67	86.33	94.83
30	96.74	96.39	96.21	95.35	99.35
36	98.78	98.64	98.55	98.21	100.25
40	99.32	99.23	99.19	98.99	100.37
48	99.74	99.71	99.71	99.62	100.34
60	99.94	99.93	99.92	99.90	100.21
120	99.98	100.00	100.00	99.98	100.03

Col. (19) See page 67.
Col. (20) See pages 66, 67.
Col. (21) A 3-months moving total of a 5 of a 5 of an 8 of a 12 of 17 weights. See page 66.
Col. (22) A 2-months moving total of a 3 of a 3 of a 4 of a 6 of an 8 of a 10 of a 12 of 5 weights. See page 65.
Col. (23) $\frac{1}{1440}$ of a 3-months moving total of a 5 of an 8 of a 12 of 15 simple weights: +2, −3, 0, 0, 0, 0, 0, +3, 0, 0, 0, 0, 0, −3, +2. See pages 71, 72, 73, 74, 75.

	(24)	(25)	(26)	(27)
Sine Curve Period (Points)	43-term Approximately 5th-degree Parabolic	Whittaker-Henderson n = 3	Whittaker-Henderson n = 4	Whittaker-Henderson n = 5
12	0	31.87	11.75	4.53
15	38.78	63.52	33.13	15.00
18	69.76	83.68	59.34	34.21
20	82.05	90.56	73.19	49.31
24	94.17	96.60	89.00	74.25
30	99.39	99.08	96.84	91.62
36	100.36	99.69	98.92	97.02
40	100.44	99.83	99.42	98.39
48	100.33	99.94	99.81	99.45
60	100.18	99.99	99.95	99.86
120	100.02	100.00	100.00	100.00

Col. (24) $\frac{1}{9600}$ of a 5-months moving total of a 5 of an 8 of a 12 of 17 simple weights: $+7, -10, 0, 0, 0, 0, 0, 0, +10, 0, 0, 0, 0, 0, 0, -10, +7$. See pages 73, 74, 75.

Col. (25) See Appendix VI.

Col. (26) See Appendix VI.

Col. (27) See Appendix VI.

APPENDIX VIII

FOURTEEN DIFFERENT GRADUATIONS OF CALL MONEY RATES ON THE NEW YORK STOCK EXCHANGE FOR 97 MONTHS—JANUARY 1886 TO JANUARY 1894 INCLUSIVE.

(The various graduations have the same column numbers in Appendices IV, VII and VIII and in Chart I.)

The reader will notice, when examining this Appendix, that the various graduations have been applied to the *logarithms* of the monthly Call Money Rates. In any graduation, the first problem which presents itself is to decide what function of the variable shall be used as raw data for purposes of graduation. This problem cannot be solved by refusing to think about it. There is nothing magical in the form in which the data happen to be originally presented. For example, if the investigator were interested in the history of bond prices and bond yields, it would make an appreciable difference whether he selected prices or yields as the variable to which he would apply a graduation formula. This would be true even if the bonds were all perpetuities—when it would seem legitimate to have averaged their prices. Many economic series are of this type—where it would seem about equally reasonable to select as the raw data for graduation a series or its reciprocals. Of course, if the logarithms of a series be taken, it becomes mathematically indifferent whether the logarithms of the series or the logarithms of its reciprocals be graduated. There are always disadvantages associated with the choice of any particular function of the data—natural numbers, logarithms, reciprocals, etc.

Some of the reasons which led us to graduate the logarithms of monthly Call Money Rates, rather than the natural numbers, are concerned with the nature of the data, while others are concerned with the nature of graduation. The nature of the data is such that the significance of changes would seem to be measured better by ratios than by differences. A change from a 3% Call Money rate to a 4% rate would seem more nearly comparable with a change from a 6% rate to an 8% rate than with a change from a 6% rate to a 7% rate. As an index of change in general money market conditions, a movement from a 3% rate to a 4% rate would seem more important than a movement from a 6% rate to a 7% rate. The nature of graduation is such as to suggest graduating the logarithms of Call Money Rates rather than the natural numbers. The distribution of deviations of the rates from the graduation is more symmetrical when the graduation has been applied to the logarithms than it is when it has been applied to the natural numbers. The Call Money data when charted in the form of natural numbers tend to show flat minimum areas and sharply cusped maximum areas. If the logarithms are charted, there is a tendency for the data to show more of a sine-like appearance with the shapes of maximum areas more nearly the same as those of minimum areas. A graduation applied to the natural numbers will not give as close a fit to the cusped maximum areas as to the flat minimum areas. If the graduation be applied to the logarithms of the data, the closeness of fit will tend to be more nearly the same for maximum and minimum areas.

No function of the data can be chosen such that its graduation will not have peculiarities which, for particular purposes, might be undesirable. For example, a graduation of the logarithms of Call Money Rates will be such that if a borrower had a loan of constant size throughout a period of some years, his interest charges would be somewhat less if he paid the graduated rates than they would be if he paid the actual rates. In the case of most economic data, it is extremely difficult to be sure that some particular function of the data is overwhelmingly more significant than any other function. We decided that the logarithm was the most significant function for our purposes.

In all but two of the graduations below, data outside the range January 1886 to January 1894 have been used. The data (logarithms) for two years before 1886 and two years after 1894 are:

	1884	1885	1894	1895
J	.279	.076	.009	.130
F	.274	.158	.000	.176
M	.243	.117	.037	.352
A	.322	.130	.053	.352
M	1.176	.158	.041	.121
J	.537	.076	.000	.064
J	.279	.130	.000	.146
A	.243	.176	.000	.013
S	.243	.190	.000	.193
O	.290	.328	.000	.336
N	.158	.430	.013	.294
D	.176	.439	.158	.659

	Data (Logarithms)	(2) 2 of a 12 Simple Moving Average	(11) Spencer's 21-term 3rd-degree Parabolic	(12) Kenchington's 27-term 3rd-degree Parabolic	(13) 29-term Approximately 3rd-degree Parabolic	(14) 29-term Non-Parabolic
1886						
J	.328	.3525	.3682	.3501	.3503	.3467
F	.314	.3846	.3827	.3845	.3878	.3856
M	.423	.4317	.3983	.4220	.4264	.4266
A	.377	.4716	.4220	.4613	.4653	.4672
M	.459	.5006	.4587	.4986	.5029	.5059
J	.525	.5349	.5076	.5350	.5383	.5416
J	.352	.5680	.5650	.5709	.5725	.5748
A	.725	.5901	.6216	.6052	.6059	.6070
S	.771	.6117	.6714	.6381	.6380	.6383
O	.704	.6404	.7078	.6704	.6679	.6691
N	.751	.6680	.7288	.6987	.6950	.6981
D	.940	.6923	.7361	.7217	.7180	.7237
1887						
J	.622	.7196	.7366	.7386	.7357	.7435
F	.551	.7326	.7325	.7480	.7479	.7565
M	.703	.7295	.7288	.7509	.7552	.7626
A	.787	.7236	.7275	.7502	.7574	.7627
M	.710	.7165	.7296	.7448	.7542	.7575
J	.857	.7028	.7315	.7364	.7456	.7485
J	.677	.6908	.7308	.7263	.7308	.7348
A	.712	.6839	.7230	.7119	.7089	.7153
S	.710	.6679	.7075	.6896	.6801	.6885
O	.622	.6416	.6829	.6593	.6451	.6535
N	.663	.6074	.6476	.6187	.6044	.6106
D	.699	.5600	.6006	.5691	.5593	.5614

	Data (Logarithms)	(2) 2 of a 12 Simple Moving Average	(11) Spencer's 21-Term 3rd-Degree Parabolic	(12) Kenchington's 27-Term 3rd-Degree Parabolic	(13) 29-Term Approximately 3rd Degree Parabolic	(14) 29-Term Non-Parabolic
1888						
J	.576	.5100	.5441	.5156	.5123	.5097
F	.431	.4667	.4789	.4627	.4659	.4594
M	.439	.4340	.4097	.4148	.4230	.4146
A	.420	.4145	.3459	.3769	.3868	.3784
M	.255	.3953	.2974	.3493	.3594	.3523
J	.176	.3812	.2710	.3316	.3420	.3356
J	.158	.3754	.2700	.3251	.3332	.3274
A	.190	.3702	.2926	.3282	.3324	.3260
S	.449	.3694	.3320	.3380	.3377	.3309
O	.415	.3784	.3791	.3545	.3487	.3416
N	.408	.3918	.4232	.3761	.3659	.3591
D	.616	.4106	.4560	.4005	.3896	.3836
1889						
J	.519	.4395	.4752	.4283	.4195	.4147
F	.364	.4725	.4824	.4593	.4550	.4518
M	.486	.4988	.4831	.4920	.4942	.4927
A	.589	.5295	.4877	.5270	.5337	.5343
M	.407	.5690	.5061	.5628	.5710	.5736
J	.477	.5995	.5418	.5967	.6038	.6080
J	.550	.6268	.5941	.6285	.6313	.6364
A	.589	.6531	.6557	.6579	.6550	.6594
S	.682	.6700	.7164	.6832	.6768	.6793
O	.919	.6778	.7643	.7054	.6982	.6986
N	.853	.6913	.7910	.7250	.7189	.7179
D	.903	.7113	.7935	.7419	.7385	.7382
1890						
J	.886	.7244	.7771	.7555	.7562	.7581
F	.628	.7490	.7505	.7668	.7702	.7754
M	.628	.7750	.7255	.7745	.7789	.7873
A	.633	.7719	.7123	.7767	.7822	.7915
M	.688	.7624	.7172	.7729	.7795	.7873
J	.677	.7536	.7371	.7651	.7717	.7763
J	.663	.7328	.7626	.7549	.7602	.7610
A	1.066	.7135	.7818	.7436	.7446	.7439
S	.829	.6994	.7863	.7308	.7255	.7251
O	.699	.6876	.7707	.7145	.7039	.7054
N	.845	.6809	.7348	.6938	.6814	.6850
D	.699	.6720	.6818	.6687	.6590	.6631

	Data (Logarithms)	(2) 2 of a 12 Simple Moving Average	(11) Spencer's 21-term 3rd-degree Parabolic	(12) Kenchington's 27-term 3rd-degree Parabolic	(13) 29-term Approximately 3rd-degree Parabolic	(14) 29-term Non-parabolic
1891						
J	.591	.6518	.6288	.6390	.6365	.6397
F	.459	.6077	.5758	.6065	.6140	.6146
M	.459	.5696	.5321	.5764	.5913	.5889
A	.519	.5593	.5027	.5516	.5684	.5645
M	.641	.5478	.4894	.5324	.5466	.5430
J	.512	.5297	.4914	.5184	.5266	.5249
J	.342	.5113	.5022	.5083	.5073	.5085
A	.328	.4959	.5148	.4980	.4879	.4909
S	.653	.4828	.5220	.4843	.4669	.4701
O	.628	.4671	.5186	.4638	.4433	.4445
N	.641	.4386	.4987	.4361	.4180	.4166
D	.468	.4040	.4611	.4046	.3938	.3891
1892						
J	.380	.3859	.4077	.3755	.3728	.3656
F	.301	.3824	.3477	.3517	.3566	.3484
M	.301	.3802	.2926	.3363	.3478	.3391
A	.301	.3838	.2578	.3324	.3470	.3392
M	.176	.3919	.2531	.3401	.3546	.3481
J	.146	.4100	.2826	.3579	.3706	.3648
J	.274	.4345	.3426	.3869	.3961	.3899
A	.312	.4511	.4222	.4259	.4327	.4247
S	.616	.4840	.5058	.4752	.4797	.4715
O	.751	.5256	.5820	.5343	.5365	.5303
N	.712	.5576	.6427	.5979	.5992	.5971
D	.833	.6068	.6858	.6591	.6590	.6644
1893						
J	.602	.6659	.7151	.7139	.7092	.7231
F	.477	.7093	.7355	.7547	.7451	.7657
M	.914	.7254	.7516	.7745	.7624	.7864
A	.688	.7081	.7655	.7714	.7593	.7817
M	.556	.6724	.7703	.7454	.7354	.7528
J	.948	.6203	.7553	.6991	.6929	.7037
J	.889	.5635	.7127	.6391	.6349	.6398
A	.740	.5190	.6385	.5691	.5647	.5673
S	.574	.4625	.5355	.4919	.4884	.4907
O	.377	.3995	.4149	.4119	.4099	.4126
N	.230	.3516	.2920	.3302	.3303	.3332
D	.064	.2907	.1825	.2473	.2526	.2529
1894						
J	.009	.2141	.0976	.1681	.1787	.1742

	Data (Logarithms)	(18) 35-term 5th-degree Parabolic	(19) 41-term 5th-degree Parabolic	(20) 43-term 5th-degree Parabolic	(21) 45-term 5th-degree Parabolic
1886					
J	.328	.3459	.3519	.3528	.3530
F	.314	.3854	.3877	.3886	.3887
M	.423	.4277	.4257	.4256	.4258
A	.377	.4687	.4640	.4635	.4635
M	.459	.5060	.5014	.5012	.5011
J	.525	.5399	.5379	.5379	.5378
J	.352	.5712	.5731	.5730	.5732
A	.725	.6021	.6066	.6068	.6069
S	.771	.6338	.6383	.6387	.6388
O	.704	.6659	.6679	.6681	.6683
N	.751	.6957	.6947	.6948	.6949
D	.940	.7210	.7179	.7179	.7178
1887					
J	.622	.7399	.7368	.7365	.7365
F	.551	.7521	.7505	.7504	.7504
M	.703	.7575	.7589	.7591	.7592
A	.787	.7586	.7620	.7624	.7625
M	.710	.7561	.7598	.7599	.7601
J	.857	.7497	.7517	.7518	.7517
J	.677	.7380	.7376	.7372	.7369
A	.712	.7204	.7166	.7155	.7150
S	.710	.6946	.6881	.6863	.6857
O	.622	.6598	.6518	.6501	.6493
N	.663	.6167	.6088	.6073	.6066
D	.699	.5665	.5609	.5598	.5595
1888					
J	.576	.5119	.5106	.5107	.5107
F	.431	.4573	.4617	.4627	.4630
M	.439	.4087	.4172	.4188	.4195
A	.420	.3696	.3801	.3818	.3827
M	.255	.3432	.3518	.3535	.3541
J	.176	.3292	.3331	.3341	.3346
J	.158	.3253	.3235	.3238	.3240
A	.190	.3274	.3222	.3221	.3221
S	.449	.3337	.3283	.3282	.3282
O	.415	.3437	.3414	.3414	.3416
N	.408	.3584	.3612	.3618	.3621
D	.616	.3804	.3876	.3887	.3889

	Data (Logarithms)	(18) 35-term 5th degree Parabolic	(19) 41-term 5th-degree Parabolic	(20) 43-term 5th-degree Parabolic	(21) 45-term 5th-degree Parabolic
1889					
J	.519	.4119	.4202	.4209	.4212
F	.364	.4517	.4573	.4574	.4575
M	.486	.4973	.4967	.4964	.4961
A	.589	.5438	.5364	.5352	.5346
M	.407	.5854	.5736	.5716	.5710
J	.477	.6180	.6067	.6048	.6041
J	.550	.6413	.6348	.6338	.6332
A	.589	.6586	.6589	.6587	.6587
S	.682	.6737	.6803	.6810	.6813
O	.919	.6906	.7002	.7018	.7022
N	.853	.7108	.7199	.7212	.7218
D	.903	.7337	.7388	.7398	.7403
1890					
J	.886	.7561	.7563	.7569	.7570
F	.628	.7757	.7710	.7709	.7707
M	.628	.7882	.7816	.7802	.7800
A	.633	.7918	.7858	.7845	.7840
M	.688	.7872	.7836	.7827	.7821
J	.677	.7767	.7754	.7747	.7745
J	.663	.7616	.7621	.7619	.7620
A	1.066	.7433	.7450	.7452	.7422
S	.829	.7229	.7251	.7253	.7254
O	.699	.7010	.7033	.7036	.7038
N	.845	.6788	.6805	.6811	.6810
D	.699	.6571	.6573	.6576	.6576
1891					
J	.591	.6359	.6340	.6340	.6343
F	.459	.6119	.6107	.6114	.6115
M	.459	.5871	.5885	.5896	.5901
A	.519	.5632	.5681	.5689	.5695
M	.641	.5424	.5491	.5494	.5499
J	.512	.5259	.5306	.5307	.5308
J	.342	.5132	.5121	.5116	.5114
A	.328	.5000	.4925	.4911	.4908
S	.653	.4816	.4708	.4690	.4685
O	.628	.4567	.4467	.4449	.4412
N	.641	.4268	.4208	.4196	.4189
D	.468	.3952	.3951	.3943	.3942

		(18)	(19)	(20)	(21)
	Data (Logarithms)	35-term 5th-degree Parabolic	41-term 5th-degree Parabolic	43-term 5th-degree Parabolic	45-term 5th-degree Parabolic
1892					
J	.380	.3674	.3714	.3712	.3710
F	.301	.3471	.3519	.3516	.3518
M	.301	.3364	.3380	.3382	.3386
A	.301	.3338	.3319	.3328	.3334
M	.176	.3387	.3355	.3370	.3379
J	.146	.3502	.3505	.3524	.3538
J	.274	.3692	.3775	.3804	.3819
A	.312	.4001	.4173	.4213	.4229
S	.616	.4484	.4697	.4737	.4752
O	.751	.5141	.5321	.5350	.5362
N	.712	.5912	.5996	.6008	.6013
D	.833	.6701	.6657	.6651	.6645
1893					
J	.602	.7386	.7235	.7207	.7193
F	.477	.7866	.7660	.7614	.7597
M	.914	.8088	.7876	.7825	.7802
A	.688	.8040	.7856	.7810	.7789
M	.556	.7726	.7602	.7568	.7553
J	.948	.7179	.7141	.7125	.7116
J	.889	.6427	.6517	.6516	.6512
A	.740	.5685	.5777	.5782	.5783
S	.574	.4874	.4962	.4969	.4972
O	.377	.4069	.4109	.4114	.4115
N	.230	.3282	.3248	.3246	.3248
D	064	.2491	.2404	.2401	.2404
1894					
J	.009	.1684	.1608	.1615	.1620

	Data (Logarithms)	(22) 45-term 5th-degree Parabolic	(23) 57-term Approximately 5th-degree Parabolic	(24) 43-term Approximately 5th-degree Parabolic	(25) Whittaker-Henderson n = 3	(27) Whittaker-Henderson n = 5
1886						
J	.328	.3477	.3491	.3528	.3076	.2743
F	.314	.3890	.3867	.3885	.3428	.3304
M	.423	.4282	.4259	.4254	.3817	.3845
A	.377	.4652	.4652	.4633	.4245	.4365
M	.459	.4994	.5031	.5013	.4711	.4862
J	.525	.5369	.5388	.5380	.5206	.5333
J	.352	.5737	.5728	.5731	.5710	.5774
A	.725	.6084	.6057	.6071	.6194	.6179
S	.771	.6405	.6374	.6391	.6618	.6541
O	.704	.6694	.6676	.6687	.6951	.6854
N	.751	.6947	.6957	.6956	.7181	.7114
D	.940	.7168	.7201	.7192	.7314	.7319
1887						
J	.622	.7354	.7391	.7381	.7374	.7469
F	.551	.7499	.7525	.7524	.7398	.7566
M	.703	.7596	.7605	.7616	.7415	.7611
A	.787	.7637	.7633	.7654	.7432	.7603
M	.710	.7615	.7609	.7634	.7440	.7540
J	.857	.7523	.7539	.7559	.7422	.7420
J	.677	.7357	.7414	.7419	.7358	.7241
A	.712	.7122	.7219	.7205	.7232	.7003
S	.710	.6811	.6944	.6913	.7029	.6708
O	.622	.6438	.6587	.6545	.6736	.6361
N	.663	.6016	.6145	.6104	.6346	.5971
D	.699	.5564	.5638	.5612	.5860	.5551
1888						
J	.576	.5105	.5104	.5101	.5295	.5117
F	.431	.4657	.4579	.4601	.4689	.4689
M	.439	.4245	.4105	.4145	.4094	.4288
A	.420	.3884	.3720	.3763	.3567	.3935
M	.255	.3590	.3443	.3474	.3163	.3649
J	.176	.3370	.3270	.3279	.2928	.3415
J	.158	.3242	.3194	.3177	.2883	.3331
A	.190	.3206	.3199	.3166	.3015	.3308
S	.449	.3263	.3265	.3233	.3279	.3370
O	.415	.3403	.3387	.3371	.3611	.3505
N	.408	.3637	.3573	.3582	.3951	.3700
D	.616	.3925	.3850	.3861	.4255	.3942

	Data (Logarithms)	(22) 45-term 5th-degree Parabolic	(23) 39-term Approximately 5th-degree Parabolic	(24) 43-term Approximately 5th-degree Parabolic	(25) Whittaker-Henderson n = 3	(27) Whittaker-Henderson n = 5
1889						
J	.519	.4238	.4159	.4194	.4504	.4219
F	.364	.4593	.4551	.4572	.4709	.4522
M	.486	.4947	.4982	.4976	.4902	.4845
A	.589	.5302	.5408	.5374	.5117	.5183
M	.407	.5649	.5798	.5745	.5385	.5530
J	.477	.5982	.6129	.6079	.5726	.5880
J	.550	.6293	.6395	.6365	.6135	.6225
A	.589	.6581	.6604	.6606	.6580	.6555
S	.682	.6842	.6787	.6821	.7010	.6858
O	.919	.7071	.6970	.7024	.7367	.7123
N	.853	.7267	.7163	.7215	.7603	.7341
D	.903	.7435	.7371	.7402	.7702	.7508
1890						
J	.886	.7574	.7571	.7580	.7683	.7625
F	.628	.7687	.7742	.7727	.7596	.7697
M	.628	.7767	.7856	.7826	.7504	.7732
A	.633	.7804	.7906	.7873	.7457	.7737
M	.688	.7790	.7879	.7855	.7476	.7716
J	.677	.7725	.7783	.7771	.7548	.7669
J	.663	.7616	.7635	.7637	.7634	.7593
A	1.066	.7461	.7451	.7463	.7679	.7484
S	.829	.7270	.7243	.7257	.7630	.7339
O	.699	.7051	.7023	.7033	.7460	.7158
N	.845	.6815	.6800	.6804	.7171	.6945
D	.699	.6578	.6569	.6565	.6789	.6707
1891						
J	.591	.6346	.6331	.6327	.6360	.6453
F	.459	.6125	.6094	.6101	.5937	.6193
M	.459	.5917	.5868	.5887	.5566	.5935
A	.519	.5718	.5658	.5684	.5274	.5684
M	.641	.5519	.5472	.5498	.5070	.5442
J	.512	.5314	.5306	.5320	.4956	.5207
J	.342	.5097	.5142	.5137	.4936	.4977
A	.328	.4865	.4967	.4938	.4955	.4749
S	.653	.4625	.4764	.4718	.4962	.4519
O	.628	.4386	.4520	.4473	.4894	.4285
N	.641	.4149	.4240	.4210	.4698	.4051
D	.468	.3923	.3954	.3943	.4339	.3827

	Data (Logarithms)	(22) 45-term 5th-degree Parabolic	(23) 39-term Approximately 5th-degree Parabolic	(24) 43-term Approximately 5th-degree Parabolic	(25) Whittaker-Henderson n = 3	(27) Whittaker-Henderson n = 5
1892						
J	.380	.3714	.3693	.3693	.3903	.3630
F	.301	.3527	.3477	.3476	.3459	.3481
M	.301	.3391	.3332	.3319	.3087	.3401
A	.301	.3334	.3267	.3245	.2861	.3409
M	.176	.3390	.3290	.3271	.2838	.3519
J	.146	.3579	.3409	.3413	.3049	.3737
J	.274	.3898	.3652	.3693	.3485	.4059
A	.312	.4337	.4037	.4112	.4096	.4471
S	.616	.4869	.4565	.4661	.4803	.4950
O	.751	.5460	.5224	.5309	.5517	.5466
N	.712	.6062	.5967	.6013	.6166	.5988
D	.833	.6625	.6711	.6707	.6712	.6486
1893						
J	.602	.7095	.7364	.7309	.7147	.6931
F	.477	.7459	.7848	.7753	.7483	.7296
M	.914	.7652	.8092	.7988	.7727	.7555
A	.688	.7639	.8057	.7977	.7866	.7682
M	.556	.7465	.7761	.7724	.7872	.7652
J	.948	.7080	.7250	.7259	.7703	.7441
J	.889	.6500	.6576	.6618	.7306	.7028
A	.740	.5816	.5791	.5849	.6642	.6397
S	.574	.5005	.4957	.5004	.5698	.5538
O	.377	.4127	.4103	.4117	.4482	.4446
N	.230	.3238	.3240	.3217	.3011	.3120
D	.064	.2383	.2386	.2341	.1299	.1560
1894						
J	.009	.1611	.1567	.1528	−.0647	−.0234