Advanced calculus

Edwin Bidwell Wilson

ADVANCED CALCULUS

A TEXT UPON SELECT PARTS OF DIFFERENTIAL CALCULUS, DIFFERENTIAL EQUATIONS, INTEGRAL CALCULUS, THEORY OF FUNCTIONS, WITH NUMEROUS EXERCISES

BY

EDWIN BIDWELL WILSON, PH.D.
PROFESSOR OF MATHEMATICS IN THE MASSACHUSETTS INSTITUTE OF TECHNOLOGY

CHAPTERS I–X

GINN AND COMPANY
BOSTON · NEW YORK · CHICAGO · LONDON

911.9

The Athenæum Press
GINN AND COMPANY · PRO-
PRIETORS · BOSTON · U.S.A.

PREFACE

It is probable that almost every teacher of advanced calculus feels the need of a text suited to present conditions and adaptable to his use. To write such a book is extremely difficult, for the attainments of students who enter a second course in calculus are different, their needs are not uniform, and the viewpoint of their teachers is no less varied. Yet in view of the cost of time and money involved in producing an Advanced Calculus, in proportion to the small number of students who will use it, it seems that few teachers can afford the luxury of having their own text; and that it consequently devolves upon an author to take as unselfish and unprejudiced a view of the subject as possible, and, so far as in him lies, to produce a book which shall have the maximum flexibility and adaptability. It was the recognition of this duty that has kept the present work in a perpetual state of growth and modification during five or six years of composition. Every attempt has been made to write in such a manner that the individual teacher may feel the minimum embarrassment in picking and choosing what seems to him best to meet the needs of any particular class.

As the aim of the book is to be a working text or laboratory manual for classroom use rather than an artistic treatise on analysis, especial attention has been given to the preparation of numerous exercises which should range all the way from those which require nothing but substitution in certain formulas to those which embody important results withheld from the text for the purpose of leaving the student some vital bits of mathematics to develop. It has been fully recognized that for the student of mathematics the work on advanced calculus falls in a period of transition, — of adolescence, — in which he must grow from close reliance upon his book to a large reliance upon himself. Moreover, as a course in advanced calculus is the *ultima Thule* of the mathematical voyages of most students of physics and engineering, it is appropriate that the text placed in the hands of those who seek that goal should by its method cultivate in them the attitude of courageous explorers, and in its extent supply not only their immediate needs, but much that may be useful for later reference and independent study.

With the large necessities of the physicist and the growing requirements of the engineer, it is inevitable that the great majority of our students of calculus should need to use their mathematics readily and vigorously rather than with hesitation and rigor. Hence, although due attention has been paid to modern questions of rigor, the chief desire has been to confirm and to extend the student's working knowledge of those great algorisms of mathematics which are naturally associated with the calculus. That the compositor should have set "vigor" where "rigor" was written, might appear more amusing were it not for the suggested antithesis that there may be many who set rigor where vigor should be.

As I have had practically no assistance with either the manuscript or the proofs, I cannot expect that so large a work shall be free from errors; I can only have faith that such errors as occur may not prove seriously troublesome. To spend upon this book so much time and energy which could have been reserved with keener pleasure for various fields of research would have been too great a sacrifice, had it not been for the hope that I might accomplish something which should be of material assistance in solving one of the most difficult problems of mathematical instruction, — that of advanced calculus.

EDWIN BIDWELL WILSON

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

CONTENTS

INTRODUCTORY REVIEW

CHAPTER I

REVIEW OF FUNDAMENTAL RULES

CHAPTER II

REVIEW OF FUNDAMENTAL THEORY

PART I. DIFFERENTIAL CALCULUS

CHAPTER III

TAYLOR'S FORMULA AND ALLIED TOPICS

CHAPTER IV

PARTIAL DIFFERENTIATION; EXPLICIT FUNCTIONS

CHAPTER V

PARTIAL DIFFERENTIATION; IMPLICIT FUNCTIONS

CHAPTER VI

COMPLEX NUMBERS AND VECTORS

PART II. DIFFERENTIAL EQUATIONS

CHAPTER VII

GENERAL INTRODUCTION TO DIFFERENTIAL EQUATIONS

CHAPTER VIII

THE COMMONER ORDINARY DIFFERENTIAL EQUATIONS

CHAPTER IX

ADDITIONAL TYPES OF ORDINARY EQUATIONS

CHAPTER X

DIFFERENTIAL EQUATIONS IN MORE THAN TWO VARIABLES

ADVANCED CALCULUS

INTRODUCTORY REVIEW

CHAPTER I

REVIEW OF FUNDAMENTAL RULES

1. On differentiation. If the function $f(x)$ is interpreted as the curve $y=f(x)$,* the quotient of the increments Δy and Δx of the dependent and independent variables measured from (x_0, y_0) is

$$\frac{y-y_0}{x-x_0}=\frac{\Delta y}{\Delta x}=\frac{\Delta f(x)}{\Delta x}=\frac{f(x_0+\Delta x)-f(x_0)}{\Delta x}, \tag{1}$$

and represents the *slope of the secant* through the points $P(x_0, y_0)$ and $P'(x_0+\Delta x, y_0+\Delta y)$ on the curve. The limit approached by the quotient $\Delta y/\Delta x$ when P remains fixed and $\Delta x \doteq 0$ is the *slope of the tangent* to the curve at the point P. *This limit,*

$$\lim_{\Delta x \doteq 0}\frac{\Delta y}{\Delta x}=\lim_{\Delta x \doteq 0}\frac{f(x_0+\Delta x)-f(x_0)}{\Delta x}=f'(x_0), \tag{2}$$

is called the *derivative* of $f(x)$ for the value $x=x_0$. As the derivative may be computed for different points of the curve, it is customary to speak of the derivative as itself a function of x and write

$$\lim_{\Delta x \doteq 0}\frac{\Delta y}{\Delta x}=\lim_{\Delta x \doteq 0}\frac{f(x+\Delta x)-f(x)}{\Delta x}=f'(x). \tag{3}$$

There are numerous notations for the derivative, for instance

$$f'(x)=\frac{df(x)}{dx}=\frac{dy}{dx}=D_xf=D_xy=y'=Df=Dy.$$

* Here and throughout the work, where figures are not given, the reader should draw graphs to illustrate the statements. Training in making one's own illustrations, whether graphical or analytic, is of great value.

The first five show distinctly that the independent variable is x, whereas the last three do not explicitly indicate the variable and should not be used unless there is no chance of a misunderstanding.

2. The fundamental formulas of differential calculus are derived directly from the application of the definition (2) or (3) and from a few fundamental propositions in limits. First may be mentioned

$$\frac{dz}{dx}=\frac{dz}{dy}\frac{dy}{dx}, \text{ where } z=\phi(y) \text{ and } y=f(x). \tag{4}$$

$$\frac{dx}{dy}=\frac{df^{-1}(y)}{dy}=\frac{1}{\frac{df(x)}{dx}}=\frac{1}{\frac{dy}{dx}}. \tag{5}$$

$$D(u\pm v)=Du\pm Dv, \qquad D(uv)=uDv+vDu. \tag{6}$$

$$D\left(\frac{u}{v}\right)=\frac{vDu-uDv}{v^2},^* \qquad D(x^n)=nx^{n-1}. \tag{7}$$

It may be recalled that (4), which is the rule for differentiating a function of a function, follows from the application of the theorem that the limit of a product is the product of the limits to the fractional identity $\frac{\Delta z}{\Delta x}=\frac{\Delta z}{\Delta y}\frac{\Delta y}{\Delta x}$; whence

$$\lim_{\Delta x\doteq 0}\frac{\Delta z}{\Delta x}=\lim_{\Delta x\doteq 0}\frac{\Delta z}{\Delta y}\cdot\lim_{\Delta x\doteq 0}\frac{\Delta y}{\Delta x}=\lim_{\Delta y\doteq 0}{}^{\dagger}\frac{\Delta z}{\Delta y}\cdot\lim_{\Delta x\doteq 0}\frac{\Delta y}{\Delta x},$$

which is equivalent to (4). Similarly, if $y=f(x)$ and if x, as the inverse function of y, be written $x=f^{-1}(y)$ from analogy with $y=\sin x$ and $x=\sin^{-1}y$, the relation (5) follows from the fact that $\Delta x/\Delta y$ and $\Delta y/\Delta x$ are reciprocals. The next three result from the immediate application of the theorems concerning limits of sums, products, and quotients (§ 21). The rule for differentiating a power is derived in case n is integral by the application of the binomial theorem.

$$\frac{\Delta y}{\Delta x}=\frac{(x+\Delta x)^n-x^n}{\Delta x}=nx^{n-1}+\frac{n(n-1)}{2\,!}x^{n-2}\Delta x+\cdots+(\Delta x)^{n-1},$$

and the limit when $\Delta x\doteq 0$ is clearly nx^{n-1}. The result may be extended to rational values of the index n by writing $n=\frac{p}{q}$, $y=x^{\frac{p}{q}}$, $y^q=x^p$ and by differentiating both sides of the equation and reducing. To prove that (7) still holds when n is irrational, it would be *necessary to have a workable definition of irrational numbers* and to develop the properties of such numbers in greater detail than seems wise at this point. The formula is therefore assumed in accordance with *the principle of permanence of form* (§ 178), just as formulas like $a^m a^n=a^{m+n}$ of the theory of exponents, which may readily be proved for rational bases and exponents, are assumed without proof to hold also for irrational bases and exponents. See, however, §§ 18–25 and the exercises thereunder.

* It is frequently better to regard the quotient as the product $u\cdot v^{-1}$ and apply (6).
† For when $\Delta x\doteq 0$, then $\Delta y\doteq 0$ or $\Delta y/\Delta x$ could not approach a limit.

3. Second may be mentioned the formulas for the derivatives of the trigonometric and the inverse trigonometric functions.

$$D\sin x = \cos x, \qquad D\cos x = -\sin x, \tag{8}$$

or

$$D\sin x = \sin(x + \tfrac{1}{2}\pi), \qquad D\cos x = \cos(x + \tfrac{1}{2}\pi), \tag{8'}$$

$$D\tan x = \sec^2 x, \qquad D\cot x = -\csc^2 x, \tag{9}$$

$$D\sec x = \sec x \tan x, \qquad D\csc x = -\csc x \cot x, \tag{10}$$

$$D\operatorname{vers} x = \sin x, \quad \text{where} \quad \operatorname{vers} x = 1 - \cos x = 2\sin^2 \tfrac{1}{2}x, \tag{11}$$

$$D\sin^{-1}x = \frac{\pm 1}{\sqrt{1-x^2}}, \qquad \begin{cases} + \text{ in quadrants I, IV,} \\ - \text{ `` `` II, III,} \end{cases} \tag{12}$$

$$D\cos^{-1}x = \frac{\pm 1}{\sqrt{1-x^2}}, \qquad \begin{cases} - \text{ in quadrants I, II,} \\ + \text{ `` `` III, IV,} \end{cases} \tag{13}$$

$$D\tan^{-1}x = \frac{1}{1+x^2}, \qquad D\cot^{-1}x = -\frac{1}{1+x^2}, \tag{14}$$

$$D\sec^{-1}x = \frac{\pm 1}{x\sqrt{x^2-1}}, \qquad \begin{cases} + \text{ in quadrants I, III,} \\ - \text{ `` `` II, IV,} \end{cases} \tag{15}$$

$$D\csc^{-1}x = \frac{\pm 1}{x\sqrt{x^2-1}}, \qquad \begin{cases} - \text{ in quadrants I, III,} \\ + \text{ `` `` II, IV,} \end{cases} \tag{16}$$

$$D\operatorname{vers}^{-1}x = \frac{\pm 1}{\sqrt{2x - x^2}}, \qquad \begin{cases} + \text{ in quadrants I, II,} \\ - \text{ `` `` III, IV.} \end{cases} \tag{17}$$

It may be recalled that to differentiate $\sin x$ the definition is applied. Then

$$\frac{\Delta \sin x}{\Delta x} = \frac{\sin(x + \Delta x) - \sin x}{\Delta x} = \frac{\sin \Delta x}{\Delta x}\cos x - \frac{1 - \cos \Delta x}{\Delta x}\sin x.$$

It now is merely a question of evaluating the two limits which thus arise, namely,

$$\lim_{\Delta x \doteq 0} \frac{\sin \Delta x}{\Delta x} \quad \text{and} \quad \lim_{\Delta x \doteq 0} \frac{1 - \cos \Delta x}{\Delta x}. \tag{18}$$

From the properties of the circle it follows that these are respectively 1 and 0. Hence the derivative of $\sin x$ is $\cos x$. The derivative of $\cos x$ may be found in like manner or from the identity $\cos x = \sin(\frac{1}{2}\pi - x)$. The results for all the other trigonometric functions are derived by expressing the functions in terms of $\sin x$ and $\cos x$. And to treat the inverse functions, it is sufficient to recall the general method in (5). Thus

$$\text{if} \quad y = \sin^{-1}x, \qquad \text{then} \quad \sin y = x.$$

Differentiate both sides of the latter equation and note that $\cos y = \pm\sqrt{1 - \sin^2 y} = \pm\sqrt{1 - x^2}$ and the result for $D\sin^{-1}x$ is immediate. To ascertain which sign to use with the radical, it is sufficient to note that $\pm\sqrt{1-x^2}$ is $\cos y$, which is positive when the angle $y = \sin^{-1}x$ is in quadrants I and IV, negative in II and III. Similarly for the other inverse functions.

EXERCISES *

1. Carry through the derivation of (7) when $n = p/q$, and review the proofs of typical formulas selected from the list (5)–(17). Note that the formulas are often given as $D_x u^n = nu^{n-1} D_x u$, $D_x \sin u = \cos u \, D_x u$, $\cdots$, and may be derived in this form directly from the definition (3).

2. Derive the two limits necessary for the differentiation of $\sin x$.

3. Draw graphs of the inverse trigonometric functions and label the portions of the curves which correspond to quadrants I, II, III, IV. Verify the sign in (12)–(17) from the slope of the curves.

4. Find $D \tan x$ and $D \cot x$ by applying the definition (3) directly.

5. Find $D \sin x$ by the identity $\sin u - \sin v = 2 \cos \frac{u+v}{2} \sin \frac{u-v}{2}$.

6. Find $D \tan^{-1} x$ by the identity $\tan^{-1} u - \tan^{-1} v = \tan^{-1} \frac{u-v}{1+uv}$ and (3).

7. Differentiate the following expressions:

(α) $\csc 2x - \cot 2x$, (β) $\frac{1}{3} \tan^3 x - \tan x + x$, ($\gamma$) $x \cos^{-1} x - \sqrt{1-x^2}$,

(δ) $\sec^{-1} \frac{1}{\sqrt{1-x^2}}$, ($\epsilon$) $\sin^{-1} \frac{x}{\sqrt{1+x^2}}$, ($\zeta$) $x\sqrt{a^2 - x^2} + a^2 \sin^{-1} \frac{x}{a}$,

(η) $a \operatorname{vers}^{-1} \frac{x}{a} - \sqrt{2ax - x^2}$, ($\theta$) $\cot^{-1} \frac{2ax}{x^2 - a^2} - 2 \tan^{-1} \frac{x}{a}$.

What trigonometric identities are suggested by the answers for the following:

(α) $\sec^2 x$, (δ) $\frac{1}{\sqrt{1-x^2}}$, (ϵ) $\frac{1}{1+x^2}$, (θ) 0?

8. In B. O. Peirce's "Short Table of Integrals" (revised edition) differentiate the right-hand members to confirm the formulas: Nos. 31, 45–47, 91–97, 125, 127–128, 131–135, 161–163, 214–216, 220, 260–269, 294–298, 300, 380–381, 386–394.

9. If x is measured in degrees, what is $D \sin x$?

4. The logarithmic, exponential, and hyperbolic functions. The next set of formulas to be cited are

$$D \log_e x = \frac{1}{x}, \qquad D \log_a x = \frac{\log_a e}{x}, \tag{19}$$

$$De^x = e^x, \qquad Da^x = a^x \log_e a.\dagger \tag{20}$$

It may be recalled that the procedure for differentiating the logarithm is

$$\frac{\Delta \log_a x}{\Delta x} = \frac{\log_a (x + \Delta x) - \log_a x}{\Delta x} = \frac{1}{\Delta x} \log_a \frac{x + \Delta x}{x} = \frac{1}{x} \log_a \left(1 + \frac{\Delta x}{x}\right)^{\frac{x}{\Delta x}}.$$

* The student should keep on file his solutions of at least the important exercises; many subsequent exercises and considerable portions of the text depend on previous exercises.

† As is customary, the subscript e will hereafter be omitted and the symbol log will denote the logarithm to the base e; any base other than e must be specially designated as such. This observation is particularly necessary with reference to the common base 10 used in computation.

If now $x/\Delta x$ be set equal to h, the problem becomes that of evaluating

$$\lim_{h=\infty}\left(1+\frac{1}{h}\right)^h = e = 2.71828\cdots,^* \qquad \log_{10} e = 0.434294\cdots; \qquad (21)$$

and hence if e be chosen as the base of the system, $D\log x$ takes the simple form $1/x$. The exponential functions e^x and a^x may be regarded as the inverse functions of $\log x$ and $\log_a x$ in deducing (21). Further it should be noted that it is frequently useful to take the logarithm of an expression before differentiating. This is known as *logarithmic differentiation* and is used for products and complicated powers and roots. Thus

$$\text{if} \quad y = x^x, \qquad \text{then} \quad \log y = x \log x,$$

$$\text{and} \quad \frac{1}{y} y' = 1 + \log x \qquad \text{or} \quad y' = x^x(1+\log x).$$

It is the expression y'/y which is called the *logarithmic derivative* of y. An especially noteworthy property of the function $y = Ce^x$ is that the function and its derivative are equal, $y' = y$; and more generally *the function $y = Ce^{kx}$ is proportional to its derivative, $y' = ky$.*

5. The *hyperbolic functions* are the hyperbolic sine and cosine,

$$\sinh x = \frac{e^x - e^{-x}}{2}, \qquad \cosh x = \frac{e^x + e^{-x}}{2}; \qquad (22)$$

and the related functions $\tanh x$, $\coth x$, $\operatorname{sech} x$, $\operatorname{csch} x$, derived from them by the same ratios as those by which the corresponding trigonometric functions are derived from $\sin x$ and $\cos x$. From these definitions in terms of exponentials follow the formulas:

$$\cosh^2 x - \sinh^2 x = 1, \qquad \tanh^2 x + \operatorname{sech}^2 x = 1, \qquad (23)$$

$$\sinh(x \pm y) = \sinh x \cosh y \pm \cosh x \sinh y, \qquad (24)$$

$$\cosh(x \pm y) = \cosh x \cosh y \pm \sinh x \sinh y, \qquad (25)$$

$$\cosh\frac{x}{2} = +\sqrt{\frac{\cosh x + 1}{2}}, \qquad \sinh\frac{x}{2} = \pm\sqrt{\frac{\cosh x - 1}{2}}, \qquad (26)$$

$$D\sinh x = \cosh x, \qquad D\cosh x = \sinh x, \qquad (27)$$

$$D\tanh x = \operatorname{sech}^2 x, \qquad D\coth x = -\operatorname{csch}^2 x, \qquad (28)$$

$$D\operatorname{sech} x = -\operatorname{sech} x \tanh x, \quad D\operatorname{csch} x = -\operatorname{csch} x \coth x. \quad (29)$$

The inverse functions are expressible in terms of logarithms. Thus

$$y = \sinh^{-1} x, \qquad x = \sinh y = \frac{e^{2y}-1}{2e^y},$$

$$e^{2y} - 2xe^y - 1 = 0, \qquad e^y = x \pm \sqrt{x^2+1}.$$

* The treatment of this limit is far from complete in the majority of texts. Reference for a careful presentation may, however, be made to Granville's "Calculus," pp. 31–34, and Osgood's "Calculus," pp. 78–82. See also Ex. 1, (β), in § 165 below.

Here only the positive sign is available, for e^y is never negative. Hence

$$\sinh^{-1} x = \log\left(x + \sqrt{x^2+1}\right), \qquad \text{any } x, \tag{30}$$

$$\cosh^{-1} x = \log\left(x \pm \sqrt{x^2-1}\right), \qquad x^2 > 1, \tag{31}$$

$$\tanh^{-1} x = \frac{1}{2}\log\frac{1+x}{1-x}, \qquad x^2 < 1, \tag{32}$$

$$\coth^{-1} x = \frac{1}{2}\log\frac{x+1}{x-1}, \qquad x^2 > 1, \tag{33}$$

$$\operatorname{sech}^{-1} x = \log\left(\frac{1}{x} \pm \sqrt{\frac{1}{x^2} - 1}\right), \qquad x^2 < 1, \tag{34}$$

$$\operatorname{csch}^{-1} x = \log\left(\frac{1}{x} + \sqrt{\frac{1}{x^2} + 1}\right), \qquad \text{any } x, \tag{35}$$

$$D\sinh^{-1} x = \frac{+1}{\sqrt{x^2+1}}, \qquad D\cosh^{-1} x = \frac{\pm 1}{\sqrt{x^2-1}}, \tag{36}$$

$$D\tanh^{-1} x = \frac{1}{1-x^2} = D\coth^{-1} x = \frac{1}{1-x^2}, \tag{37}$$

$$D\operatorname{sech}^{-1} x = \frac{\pm 1}{x\sqrt{1-x^2}}, \qquad D\operatorname{csch}^{-1} x = \frac{-1}{x\sqrt{1+x^2}}. \tag{38}$$

EXERCISES

1. Show by logarithmic differentiation that

$$D(uvw\cdots) = \left(\frac{u'}{u} + \frac{v'}{v} + \frac{w'}{w} + \cdots\right)(uvw\cdots),$$

and hence derive the rule: To differentiate a product differentiate each factor alone and add all the results thus obtained.

2. Sketch the graphs of the hyperbolic functions, interpret the graphs as those of the inverse functions, and verify the range of values assigned to x in (30)–(35).

3. Prove sundry of formulas (23)–(29) from the definitions (22).

4. Prove sundry of (30)–(38), checking the signs with care. In cases where double signs remain, state when each applies. Note that in (31) and (34) *the double sign may be placed before the* log *for the reason that the two expressions are reciprocals.*

5. Derive a formula for $\sinh u \pm \sinh v$ by applying (24); find a formula for $\tanh \frac{1}{2}x$ analogous to the trigonometric formula $\tan \frac{1}{2}x = \sin x/(1+\cos x)$.

6. *The gudermannian.* The function $\phi = \operatorname{gd} x$, defined by the relations

$$\sinh x = \tan\phi, \quad \phi = \operatorname{gd} x = \tan^{-1}\sinh x, \quad -\tfrac{1}{2}\pi < \phi < +\tfrac{1}{2}\pi,$$

is called the gudermannian of x. Prove the set of formulas:

$$\cosh x = \sec\phi, \quad \tanh x = \sin\phi, \quad \operatorname{csch} x = \cot\phi, \quad \text{etc.};$$

$$D\operatorname{gd} x = \operatorname{sech} x, \quad x = \operatorname{gd}^{-1}\phi = \log\tan\left(\tfrac{1}{2}\phi + \tfrac{1}{4}\pi\right), \quad D\operatorname{gd}^{-1}\phi = \sec\phi.$$

7. Substitute the functions of ϕ in Ex. 6 for their hyperbolic equivalents in (23), (26), (27), and reduce to simple known trigonometric formulas.

8. Differentiate the following expressions:

(α) $(x+1)^2(x+2)^{-3}(x+3)^{-4}$, ($\beta$) $x^{\log x}$, (γ) $\log_x(x+1)$,

(δ) $x+\log\cos(x-\frac{1}{4}\pi)$, ($\epsilon$) $2\tan^{-1}e^x$, (ζ) $x-\tanh x$,

(η) $x\tanh^{-1}x+\frac{1}{2}\log(1-x^2)$, ($\theta$) $\dfrac{e^{ax}(a\sin mx-m\cos mx)}{m^2+a^2}$.

9. Check sundry formulas of Peirce's "Table," pp. 1–61, 81–82.

6. Geometric properties of the derivative. As the quotient (1) and its limit (2) give the slope of a secant and of the tangent, it appears from graphical considerations that when the derivative is positive the function is increasing with x, but decreasing when the derivative is negative.* Hence *to determine the regions in which a function is increasing or decreasing, one may find the derivative and determine the regions in which it is positive or negative.*

One must, however, be careful not to apply this rule too blindly; for in so simple a case as $f(x)=\log x$ it is seen that $f'(x)=1/x$ is positive when $x>0$ and negative when $x<0$, and yet $\log x$ has no graph when $x<0$ and is not considered as decreasing. Thus the formal derivative may be real when the function is not real, and it is therefore best to make a rough sketch of the function to corroborate the evidence furnished by the examination of $f'(x)$.

If x_0 is a value of x such that immediately† upon one side of $x=x_0$ the function $f(x)$ is increasing whereas immediately upon the other side it is decreasing, the ordinate $y_0=f(x_0)$ will be a maximum or minimum or $f(x)$ will become positively or negatively infinite at x_0. If the case where $f(x)$ becomes infinite be ruled out, one may say that *the function will have a minimum or maximum at x_0 according as the derivative changes from negative to positive or from positive to negative when x, moving in the positive direction, passes through the value x_0.* Hence *the usual rule for determining maxima and minima is to find the roots of $f'(x)=0$.*

This rule, again, must not be applied blindly. For first, $f'(x)$ may vanish where there is no maximum or minimum as in the case $y=x^3$ at $x=0$ where the derivative does not change sign; or second, $f'(x)$ may change sign by becoming infinite as in the case $y=x^{\frac{2}{3}}$ at $x=0$ where the curve has a vertical cusp, point down, and a minimum; or third, the function $f(x)$ may be restricted to a given range of values $a\leqq x\leqq b$ for x and then the values $f(a)$ and $f(b)$ of the function at the ends of the interval will in general be maxima or minima without implying that the derivative vanish. Thus although the derivative is highly useful in determining maxima and minima, it should not be trusted to the complete exclusion of the corroborative evidence furnished by a rough sketch of the curve $y=f(x)$.

* The construction of illustrative figures is again left to the reader.

† The word "immediately" is necessary because the maxima or minima may be merely *relative*; in the case of several maxima and minima in an interval, some of the maxima may actually be less than some of the minima.

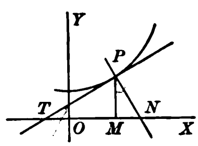
7. The derivative may be used to express the *equations of the tangent and normal*, the *values of the subtangent and subnormal*, and so on.

Equation of tangent, $y - y_0 = y'_0(x - x_0)$, (39)

Equation of normal, $(y - y_0)y'_0 + (x - x_0) = 0$, (40)

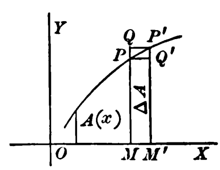
$$TM = \text{subtangent} = y_0/y'_0, \quad MN = \text{subnormal} = y_0 y'_0, \tag{41}$$

$$OT = x\text{-intercept of tangent} = x_0 - y_0/y'_0, \text{ etc.} \tag{42}$$



The derivation of these results is sufficiently evident from the figure. It may be noted that the subtangent, subnormal, etc., are numerical values for a given point of the curve but may be regarded as functions of x like the derivative.

In geometrical and physical problems it is frequently necessary to apply the definition of the derivative to finding the derivative of an unknown function. For instance if A denote the area under a curve and measured from a fixed ordinate to a variable ordinate, A is surely a function $A(x)$ of the abscissa x of the variable ordinate. If the curve is rising, as in the figure, then



$$MPQ'M' < \Delta A < MQP'M', \text{ or } y\Delta x < \Delta A < (y + \Delta y)\Delta x.$$

Divide by Δx and take the limit when $\Delta x \doteq 0$. There results

$$\lim_{\Delta x \doteq 0} y \leqq \lim_{\Delta x \doteq 0} \frac{\Delta A}{\Delta x} \leqq \lim_{\Delta x \doteq 0} (y + \Delta y).$$

Hence
$$\lim_{\Delta x \doteq 0} \frac{\Delta A}{\Delta x} = \frac{dA}{dx} = y. \tag{43}$$

Rolle's Theorem and the *Theorem of the Mean* are two important theorems on derivatives which will be treated in the next chapter but may here be stated as evident from their geometric interpretation. Rolle's Theorem states that: *If a function has a derivative at every*

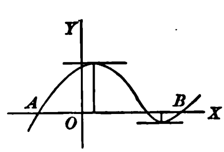


FIG. 1

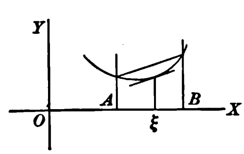


FIG. 2

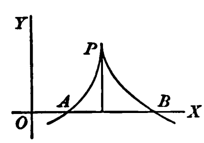


FIG. 3

point of an interval and if the function vanishes at the ends of the interval, then there is at least one point within the interval at which the derivative vanishes. This is illustrated in Fig. 1, in which there are two such points. The Theorem of the Mean states that: *If a function* *has a derivative at each point of an interval, there is at least one point in the interval such that the tangent to the curve $y = f(x)$ is parallel to the chord of the interval.* This is illustrated in Fig. 2 in which there is only one such point.

Again care must be exercised. In Fig. 3 the function vanishes at A and B but there is no point at which the slope of the tangent is zero. This is not an exception or contradiction to Rolle's Theorem for the reason that the function does not satisfy the conditions of the theorem. In fact at the point P, although there is a tangent to the curve, there is no derivative ; the quotient (1) formed for the point P becomes negatively infinite as $\Delta x \doteq 0$ from one side, positively infinite as $\Delta x \doteq 0$ from the other side, and therefore does not approach a definite limit as is required in the definition of a derivative. The hypothesis of the theorem is not satisfied and there is no reason that the conclusion should hold.

EXERCISES

1. Determine the regions in which the following functions are increasing or decreasing, sketch the graphs, and find the maxima and minima:

$$(\alpha)\ \tfrac{1}{3}x^3 - x^2 + 2, \qquad (\beta)\ (x+1)^{\frac{2}{3}}(x-5)^3, \qquad (\gamma)\ \log(x^2-4),$$
$$(\delta)\ (x-2)\sqrt{x-1}, \qquad (\epsilon)\ -(x+2)\sqrt{12-x^2}, \qquad (\zeta)\ x^3 + ax + b.$$

2. The ellipse is $r = \sqrt{x^2+y^2} = e(d+x)$ referred to an origin at the focus. Find the maxima and minima of the focal radius r, and state why $D_x r = 0$ does not give the solutions while $D_\phi r = 0$ does [the polar form of the ellipse being $r = k(1 - e\cos\phi)^{-1}$].

3. Take the ellipse as $x^2/a^2 + y^2/b^2 = 1$ and discuss the maxima and minima of the central radius $r = \sqrt{x^2+y^2}$. Why does $D_x r = 0$ give half the result when r is expressed as a function of x, and why will $D_\lambda r = 0$ give the whole result when $x = a\cos\lambda$, $y = b\sin\lambda$ and the ellipse is thus expressed in terms of the eccentric angle?

4. If $y = P(x)$ is a polynomial in x such that the equation $P(x) = 0$ has multiple roots, show that $P'(x) = 0$ for each multiple root. What more complete relationship can be stated and proved?

5. Show that the triple relation $27b^2 + 4a^3 \gtreqless 0$ determines completely the nature of the roots of $x^3 + ax + b = 0$, and state what corresponds to each possibility.

6. Define *the angle θ between two intersecting curves.* Show that

$$\tan\theta = [f'(x_0) - g'(x_0)] \div [1 + f'(x_0)g'(x_0)]$$

if $y = f(x)$ and $y = g(x)$ cut at the point (x_0, y_0).

7. Find the subnormal and subtangent of the three curves

$$(\alpha)\ y^2 = 4px, \qquad (\beta)\ x^2 = 4py, \qquad (\gamma)\ x^2 + y^2 = a^2.$$

8. *The pedal curve.* The locus of the foot of the perpendicular dropped from a fixed point to a variable tangent of a given curve is called the pedal of the given curve with respect to the given point. Show that if the fixed point is the origin, the pedal of $y = f(x)$ may be obtained by eliminating x_0, y_0, y'_0 from the equations

$$y - y_0 = y'_0(x - x_0), \quad yy'_0 + x = 0, \quad y_0 = f(x_0), \quad y'_0 = f'(x_0).$$

Find the pedal (α) of the hyperbola with respect to the center and (β) of the parabola with respect to the vertex and (γ) the focus. Show (δ) that the pedal of the parabola with respect to any point is a cubic.

9. If the curve $y = f(x)$ be revolved about the x-axis and if $V(x)$ denote the volume of revolution thus generated when measured from a fixed plane perpendicular to the axis out to a variable plane perpendicular to the axis, show that $D_x V = \pi y^2$.

10. More generally if $A(x)$ denote the area of the section cut from a solid by a plane perpendicular to the x-axis, show that $D_x V = A(x)$.

11. If $A(\phi)$ denote the sectorial area of a plane curve $r = f(\phi)$ and be measured from a fixed radius to a variable radius, show that $D_\phi A = \frac{1}{2} r^2$.

12. If ρ, h, p are the density, height, pressure in a vertical column of air, show that $dp/dh = -\rho$. If $\rho = kp$, show $p = Ce^{-kh}$.

13. Draw a graph to illustrate an apparent exception to the Theorem of the Mean analogous to the apparent exception to Rolle's Theorem, and discuss.

14. Show that the analytic statement of the Theorem of the Mean for $f(x)$ is that a value $x = \xi$ intermediate to a and b may be found such that

$$f(b) - f(a) = f'(\xi)(b - a), \qquad a < \xi < b.$$

15. Show that the semiaxis of an ellipse is a mean proportional between the x-intercept of the tangent and the abscissa of the point of contact.

16. Find the values of the length of the tangent (α) from the point of tangency to the x-axis, (β) to the y-axis, (γ) the total length intercepted between the axes. Consider the same problems for the normal (figure on page 8).

17. Find the angle of intersection of (α) $y^2 = 2\,mx$ and $x^2 + y^2 = a^2$, (β) $x^2 = 4\,ay$ and $y = \dfrac{8\,a^3}{x^2 + 4\,a^2}$, ($\gamma$) $\dfrac{x^2}{a^2 - \lambda^2} + \dfrac{y^2}{b^2 - \lambda^2} = 1$ for $0 < \lambda < b$ and $b < \lambda < a$.

18. A constant length is laid off along the normal to a parabola. Find the locus.

19. The length of the tangent to $x^{\frac{2}{3}} + y^{\frac{2}{3}} = a^{\frac{2}{3}}$ intercepted by the axes is constant.

20. The triangle formed by the asymptotes and any tangent to a hyperbola has constant area.

21. Find the length PT of the tangent to $x = \sqrt{c^2 - y^2} + c \operatorname{sech}^{-1}(y/c)$.

22. Find the greatest right cylinder inscribed in a given right cone.

23. Find the cylinder of greatest lateral surface inscribed in a sphere.

24. From a given circular sheet of metal cut out a sector that will form a cone (without base) of maximum volume.

25. Join two points A, B in the same side of a line to a point P of the line in such a way that the distance $PA + PB$ shall be least.

26. Obtain the formula for the distance from a point to a line as the minimum distance.

27. *Test for maximum or minimum.* (α) If $f(x)$ vanishes at the ends of an interval and is positive within the interval and if $f'(x) = 0$ has only one root in the interval, that root indicates a maximum. Prove this by Rolle's Theorem. Apply it in Exs. 22–24. (β) If $f(x)$ becomes indefinitely great at the ends of an interval and $f'(x) = 0$ has only one root in the interval, that root indicates a minimum.

Prove by Rolle's Theorem, and apply in Exs. 25–26. These rules or various modifications of them generally suffice in practical problems to distinguish between maxima and minima without examining either the changes in sign of the first derivative or the sign of the second derivative; for generally there is only one root of $f'(x) = 0$ in the region considered.

28. Show that $x^{-1}\sin x$ from $x = 0$ to $x = \frac{1}{2}\pi$ steadily decreases from 1 to $2/\pi$.

29. If $0 < x < 1$, show $(\alpha)\ 0 < x - \log(1+x) < \frac{1}{2}x^2$, $(\beta)\ \frac{\frac{1}{2}x^2}{1+x} < x - \log(1+x)$.

30. If $0 > x > -1$, show that $\frac{1}{2}x^2 < x - \log(1+x) < \frac{\frac{1}{2}x^2}{1+x}$.

8. Derivatives of higher order. The derivative of the derivative (regarded as itself a function of x) is the second derivative, and so on to the nth derivative. Customary notations are:

$$f''(x) = \frac{d^2f(x)}{dx^2} = \frac{d^2y}{dx^2} = D_x^2 f = D_x^2 y = y'' = D^2 f = D^2 y,$$

$$f'''(x),\ f^{\text{iv}}(x),\ \cdots,\ f^{(n)}(x);\quad \frac{d^3y}{dx^3},\ \frac{d^4y}{dx^4},\ \cdots,\ \frac{d^ny}{dx^n},\ \cdots.$$

The nth derivative of the sum or difference is the sum or difference of the nth derivatives. For the nth derivative of the product there is a special formula known as *Leibniz's Theorem.* It is

$$D^n(uv) = D^n u \cdot v + nD^{n-1}uDv + \frac{n(n-1)}{2!}D^{n-2}uD^2v + \cdots + uD^n v. \quad (44)$$

This result may be written in symbolic form as

$$\text{Leibniz's Theorem}\quad D^n(uv) = (Du + Dv)^n, \quad (44')$$

where it is to be understood that in expanding $(Du + Dv)^n$ the term $(Du)^k$ is to be replaced by $D^k u$ and $(Du)^0$ by $D^0 u = u$. In other words the powers refer to repeated differentiations.

A proof of (44) by induction will be found in § 27. The following proof is interesting on account of its ingenuity. Note first that from

$$D(uv) = uDv + vDu,\quad D^2(uv) = D(uDv) + D(vDu),$$

and so on, it appears that $D^2(uv)$ consists of a sum of terms, in each of which there are two differentiations, with numerical coefficients independent of u and v. In like manner it is clear that

$$D^n(uv) = C_0 D^n u \cdot v + C_1 D^{n-1}uDv + \cdots + C_{n-1}DuD^{n-1}v + C_n uD^n v$$

is a sum of terms, in each of which there are n differentiations, with coefficients C independent of u and v. To determine the C's any suitable functions u and v, say,

$$u = e^x,\quad v = e^{ax},\quad uv = e^{(1+a)x},\quad D^k e^{ax} = a^k e^{ax},$$

may be substituted. If the substitution be made and $e^{(1+a)x}$ be canceled,

$$e^{-(1+a)x}D^n(uv) = (1+a)^n = C_0 + C_1 a + \cdots + C_{n-1}a^{n-1} + C_n a^n,$$

and hence the C's are the coefficients in the binomial expansion of $(1+a)^n$.

Formula (4) for the derivative of a function of a function may be extended to higher derivatives by repeated application. More generally *any desired change of variable may be made by the repeated use of* (4) *and* (5). For if x and y be expressed in terms of known functions of new variables u and v, it is always possible to obtain the derivatives $D_x y$, $D_x^2 y$, $\cdots$ in terms of $D_u v$, $D_u^2 v$, $\cdots$, and thus any expression $F(x, y, y', y'', \cdots)$ may be changed into an equivalent expression $\Phi(u, v, v', v'', \cdots)$ in the new variables. In each case that arises the transformations should be carried out by repeated application of (4) and (5) rather than by substitution in any general formulas.

The following typical cases are illustrative of the method of change of variable. Suppose only the dependent variable y is to be changed to z defined as $y = f(z)$. Then

$$\frac{d^2y}{dx^2} = \frac{d}{dx}\left(\frac{dy}{dx}\right) = \frac{d}{dx}\left(\frac{dz}{dx}\frac{dy}{dz}\right) = \frac{d^2z}{dx^2}\frac{dy}{dz} + \frac{dz}{dx}\left(\frac{d}{dx}\frac{dy}{dz}\right)$$
$$= \frac{d^2z}{dx^2}\frac{dy}{dz} + \frac{dz}{dx}\left(\frac{d}{dz}\frac{dy}{dz}\frac{dz}{dx}\right) = \frac{d^2z}{dx^2}\frac{dy}{dz} + \left(\frac{dz}{dx}\right)^2\frac{d^2y}{dz^2}.$$

As the derivatives of $y = f(z)$ are known, the derivative d^2y/dx^2 has been expressed in terms of z and derivatives of z with respect to x. The third derivative would be found by repeating the process. If the problem were to change the independent variable x to z, defined by $x = f(z)$,

$$\frac{dy}{dx} = \frac{dy}{dz}\frac{dz}{dx} = \frac{dy}{dz}\left(\frac{dx}{dz}\right)^{-1}, \qquad \frac{d^2y}{dx^2} = \frac{d}{dx}\left[\frac{dy}{dz}\left(\frac{dx}{dz}\right)^{-1}\right].$$
$$\frac{d^2y}{dx^2} = \frac{d^2y}{dz^2}\frac{dz}{dx}\left(\frac{dx}{dz}\right)^{-1} - \frac{dy}{dz}\left(\frac{dx}{dz}\right)^{-2}\frac{dz}{dx}\frac{d^2x}{dz^2} = \left[\frac{d^2y}{dz^2}\frac{dx}{dz} - \frac{d^2x}{dz^2}\frac{dy}{dz}\right] \div \left(\frac{dx}{dz}\right)^3.$$

The change is thus made as far as derivatives of the second order are concerned. If the change of both dependent and independent variables was to be made, the work would be similar. Particularly useful changes are to find the derivatives of y by x when y and x are expressed parametrically as functions of t, or when both are expressed in terms of new variables r, ϕ as $x = r\cos\phi$, $y = r\sin\phi$. For these cases see the exercises.

9. The *concavity of a curve* $y = f(x)$ is given by the table:

if $f''(x_0) > 0$,	the curve is concave up at $x = x_0$,
if $f''(x_0) < 0$,	the curve is concave down at $x = x_0$,
if $f''(x_0) = 0$,	an inflection point at $x = x_0$. (?)

Hence the *criterion for distinguishing between maxima and minima:*

if $f'(x_0) = 0$ and $f''(x_0) > 0$,	a minimum at $x = x_0$,
if $f'(x_0) = 0$ and $f''(x_0) < 0$,	a maximum at $x = x_0$,
if $f'(x_0) = 0$ and $f''(x_0) = 0$,	neither max. nor min. (?)

The question points are necessary in the third line because the statements are not always true unless $f'''(x_0) \neq 0$ (see Ex. 7 under § 39).

It may be recalled that the reason that the curve is concave up in case $f''(x_0) > 0$ is because the derivative $f'(x)$ is then an increasing function in the neighborhood of $x = x_0$; whereas if $f''(x_0) < 0$, the derivative $f'(x)$ is a decreasing function and the curve is convex up. It should be noted that concave up is not the same as concave toward the x-axis, except when the curve is below the axis. With regard to the use of the second derivative as a criterion for distinguishing between maxima and minima, it should be stated that in practical examples the criterion is of relatively small value. It is usually shorter to discuss the change of sign of $f'(x)$ directly, — and indeed in most cases either a rough graph of $f(x)$ or the physical conditions of the problem which calls for the determination of a maximum or minimum will immediately serve to distinguish between them (see Ex. 27 above).

The second derivative is fundamental in dynamics. By definition the *average velocity* $\bar{v}$ of a particle is the ratio of the space traversed to the time consumed, $\bar{v} = s/t$. The *actual velocity* v at any time is the limit of this ratio when the interval of time is diminished and approaches zero as its limit. Thus

$$\bar{v} = \frac{\Delta s}{\Delta t} \quad \text{and} \quad v = \lim_{\Delta t \doteq 0} \frac{\Delta s}{\Delta t} = \frac{ds}{dt}. \tag{45}$$

In like manner if a particle describes a straight line, say the x-axis, the *average acceleration* $\bar{f}$ is the ratio of the increment of velocity to the increment of time, and the *actual acceleration* f at any time is the limit of this ratio as $\Delta t \doteq 0$. Thus

$$\bar{f} = \frac{\Delta v}{\Delta t} \quad \text{and} \quad f = \lim_{\Delta t \doteq 0} \frac{\Delta v}{\Delta t} = \frac{dv}{dt} = \frac{d^2x}{dt^2}. \tag{46}$$

By *Newton's Second Law of Motion, the force acting on the particle is equal to the rate of change of momentum* with the time, momentum being defined as the product of the mass and velocity. Thus

$$F = \frac{d(mv)}{dt} = m\frac{dv}{dt} = mf = m\frac{d^2x}{dt^2}, \tag{47}$$

where it has been assumed in differentiating that the mass is constant, as is usually the case. Hence (47) appears as the fundamental equation for rectilinear motion (see also §§ 79, 84). It may be noted that

$$F = mv\frac{dv}{dx} = \frac{d}{dx}\left(\frac{1}{2}mv^2\right) = \frac{dT}{dx}, \tag{47'}$$

where $T = \frac{1}{2}mv^2$ denotes by definition the *kinetic energy* of the particle. For comments see Ex. 6 following.

EXERCISES

1. State and prove the extension of Leibniz's Theorem to products of three or more factors. Write out the square and cube of a trinomial.

2. Write, by Leibniz's Theorem, the second and third derivatives:

(α) $e^x \sin x$, (β) $\cosh x \cos x$, (γ) $x^2 e^x \log x$.

3. Write the nth derivatives of the following functions, of which the last three should first be simplified by division or separation into partial fractions.

(α) $\sqrt{x+1}$, (β) $\log(ax+b)$, (γ) $(x^2+1)(x+1)^{-3}$,

(δ) $\cos ax$, (ϵ) $e^x \sin x$, (ζ) $(1-x)/(1+x)$,

(η) $\dfrac{1}{x^2-1}$, (θ) $\dfrac{x^3+x+1}{x-1}$, (ι) $\left(\dfrac{ax+1}{ax-1}\right)^2$.

4. If y and x are each functions of t, show that

$$\frac{d^2y}{dx^2} = \frac{\dfrac{dx}{dt}\dfrac{d^2y}{dt^2} - \dfrac{dy}{dt}\dfrac{d^2x}{dt^2}}{\left(\dfrac{dx}{dt}\right)^3} = \frac{x'y'' - y'x''}{x'^3},$$

$$\frac{d^3y}{dx^3} = \frac{x'(x'y''' - y'x''') - 3x''(x'y'' - y'x'')}{x'^5}.$$

5. Find the inflection points of the curve $x = 4\phi - 2\sin\phi$, $y = 4 - 2\cos\phi$.

6. Prove (47'). Hence infer that the force which is the time-derivative of the momentum mv by (47) is also the space-derivative of the kinetic energy.

7. If A denote the area under a curve, as in (43), find $dA/d\theta$ for the curves

(α) $y = a(1-\cos\theta)$, $x = a(\theta - \sin\theta)$, (β) $x = a\cos\theta$, $y = b\sin\theta$.

8. Make the indicated change of variable in the following equations:

(α) $\dfrac{d^2y}{dx^2} + \dfrac{2x}{1+x^2}\dfrac{dy}{dx} + \dfrac{y}{(1+x^2)^2} = 0$, $x = \tan z$. *Ans.* $\dfrac{d^2y}{dz^2} + y = 0$.

(β) $(1-x^2)\left[\dfrac{d^2y}{dx^2} - \dfrac{1}{y}\left(\dfrac{dy}{dx}\right)^2\right] - x\dfrac{dy}{dx} + y = 0$, $y = e^v$, $x = \sin u$.

Ans. $\dfrac{d^2v}{du^2} + 1 = 0$.

9. *Transformation to polar coördinates.* Suppose that $x = r\cos\phi$, $y = r\sin\phi$. Then

$$\frac{dx}{d\phi} = \frac{dr}{d\phi}\cos\phi - r\sin\phi, \qquad \frac{dy}{d\phi} = \frac{dr}{d\phi}\sin\phi + r\cos\phi,$$

and so on for higher derivatives. Find $\dfrac{dy}{dx}$ and $\dfrac{d^2y}{dx^2} = \dfrac{r^2 + 2(D_\phi r)^2 - rD_\phi^2 r}{(\cos\phi\, D_\phi r - r\sin\phi)^3}$.

10. Generalize formula (5) for the differentiation of an inverse function. Find d^2x/dy^2 and d^3x/dy^3. Note that these may also be found from Ex. 4.

11. A point describes a circle with constant speed. Find the velocity and acceleration of the projection of the point on any fixed diameter.

12. Prove $\dfrac{d^2y}{dx^2} = 2uv^3 + 4v^4\left(\dfrac{dv}{du}\right)^{-1} - v^5\dfrac{d^2v}{du^2}\left(\dfrac{dv}{du}\right)^{-3}$ if $x = \dfrac{1}{v}$, $y = uv$.

10. The indefinite integral. To integrate a function $f(x)$ is to *find a function* $F(x)$ *the derivative of which is* $f(x)$. The integral $F(x)$ is not uniquely determined by the integrand $f(x)$; for any two functions which differ merely by an additive constant have the same derivative. In giving formulas for integration the constant may be omitted and understood; but in applications of integration to actual problems it should always be inserted and must usually be determined to fit the requirements of special conditions imposed upon the problem and known as the *initial conditions.*

It must not be thought that the constant of integration always appears added to the function $F(x)$. It may be combined with $F(x)$ so as to be somewhat disguised. Thus

$$\log x, \quad \log x + C, \quad \log Cx, \quad \log(x/C)$$

are all integrals of $1/x$, and all except the first have the constant of integration C, although only in the second does it appear as formally additive. To illustrate the determination of the constant by initial conditions, consider the problem of finding the area under the curve $y = \cos x$. By (43)

$$D_xA = y = \cos x \quad \text{and hence} \quad A = \sin x + C.$$

If the area is to be measured from the ordinate $x = 0$, then $A = 0$ when $x = 0$, and by direct substitution it is seen that $C = 0$. Hence $A = \sin x$. But if the area be measured from $x = -\frac{1}{2}\pi$, then $A = 0$ when $x = -\frac{1}{2}\pi$ and $C = 1$. Hence $A = 1 + \sin x$. In fact the area under a curve is not definite until the ordinate from which it is measured is specified, and the constant is needed to allow the integral to fit this initial condition.

11. The fundamental formulas of integration are as follows:

$$\int \frac{1}{x} = \log x, \qquad \int x^n = \frac{1}{n+1} x^n \text{ if } n \neq -1, \tag{48}$$

$$\int e^x = e^x, \qquad \int a^x = a^x/\log a, \tag{49}$$

$$\int \sin x = -\cos x, \qquad \int \cos x = \sin x, \tag{50}$$

$$\int \tan x = -\log \cos x, \qquad \int \cot x = \log \sin x, \tag{51}$$

$$\int \sec^2 x = \tan x, \qquad \int \csc^2 x = -\cot x, \tag{52}$$

$$\int \tan x \sec x = \sec x, \qquad \int \cot x \csc x = -\csc x, \tag{53}$$

with formulas similar to (50)–(53) for the hyperbolic functions. Also

$$\int \frac{1}{1+x^2} = \tan^{-1} x \text{ or } -\cot^{-1} x, \quad \int \frac{1}{1-x^2} = \tanh^{-1} x \text{ or } \coth^{-1} x, \tag{54}$$

$$\int\frac{1}{\sqrt{1-x^2}}=\sin^{-1}x \text{ or } -\cos^{-1}x,\quad \int\frac{\pm 1}{\sqrt{1+x^2}}=\pm\sinh^{-1}x, \tag{55}$$

$$\int\frac{1}{x\sqrt{x^2-1}}=\sec^{-1}x \text{ or } -\csc^{-1}x,\quad \int\frac{\pm 1}{x\sqrt{1-x^2}}=\mp\operatorname{sech}^{-1}x, \tag{56}$$

$$\int\frac{\pm 1}{\sqrt{x^2-1}}=\pm\cosh^{-1}x,\qquad \int\frac{\pm 1}{x\sqrt{1+x^2}}=\mp\operatorname{csch}^{-1}x, \tag{57}$$

$$\int\frac{1}{\sqrt{2x-x^2}}=\operatorname{vers}^{-1}x,\qquad \int\sec x=\operatorname{gd}^{-1}x=\log\tan\left(\frac{\pi}{4}+\frac{x}{2}\right). \tag{58}$$

For the integrals expressed in terms of the inverse hyperbolic functions, the logarithmic equivalents are sometimes preferable. This is not the case, however, in the many instances in which the problem calls for immediate solution with regard to x. Thus if $y=\int(1+x^2)^{-\frac{1}{2}}=\sinh^{-1}x+C$, then $x=\sinh(y-C)$, and the solution is effected and may be translated into exponentials. This is not so easily accomplished from the form $y=\log\left(x+\sqrt{1+x^2}\right)+C$. For this reason and because the inverse hyperbolic functions are briefer and offer striking analogies with the inverse trigonometric functions, it has been thought better to use them in the text and allow the reader to make the necessary substitutions from the table (30)–(35) in case the logarithmic form is desired.

12. In addition to these special integrals, which are consequences of the corresponding formulas for differentiation, there are the general rules of integration which arise from (4) and (6).

$$\int\frac{dz}{dy}\frac{dy}{dx}=\int\frac{dz}{dx}=z, \tag{59}$$

$$\int(u+v-w)=\int u+\int v-\int w, \tag{60}$$

$$uv=\int uv'+\int u'v. \tag{61}$$

Of these rules the second needs no comment and the third will be treated later. Especial attention should be given to the first. For instance suppose it were required to integrate $2\log x/x$. This does not fall under any of the given types; but

$$\frac{2}{x}\log x=\frac{d(\log x)^2}{d\log x}\frac{d\log x}{dx}=\frac{dz}{dy}\frac{dy}{dx}.$$

Here $(\log x)^2$ takes the place of z and $\log x$ takes the place of y. The integral is therefore $(\log x)^2$ as may be verified by differentiation. In general, it may be possible to see that a given integrand is separable into two factors, of which one is integrable when considered as a function of some function of x, while the other is the derivative of that function. Then (59) applies. Other examples are:

$$\int e^{\sin x}\cos x,\qquad \int\tan^{-1}x/(1+x^2),\qquad \int x^2\sin(x^3).$$

In the first, $z = e^y$ is integrable and as $y = \sin x$, $y' = \cos x$; in the second, $z = y$ is integrable and as $y = \tan^{-1}x$, $y' = (1 + x^2)^{-1}$; in the third $z = \sin y$ is integrable and as $y = x^3$, $y' = 3x^2$. The results are

$$e^{\sin x}, \qquad \tfrac{1}{2}(\tan^{-1}x)^2, \qquad -\tfrac{1}{3}\cos(x^3).$$

This method of integration at sight covers such a large percentage of the cases that arise in geometry and physics that it must be thoroughly mastered.*

EXERCISES

1. Verify the fundamental integrals (48)–(58) and give the hyperbolic analogues of (50)–(53).

2. Tabulate the integrals here expressed in terms of inverse hyperbolic functions by means of the corresponding logarithmic equivalents.

3. Write the integrals of the following integrands at sight:

(α) $\sin ax$, (β) $\cot(ax + b)$, (γ) $\tanh 3x$,

(δ) $\dfrac{1}{a^2 + x^2}$, (ϵ) $\dfrac{1}{\sqrt{x^2 - a^2}}$, ($\zeta$) $\dfrac{1}{\sqrt{2ax - x^2}}$,

(η) $\dfrac{1}{x \log x}$, (θ) $\dfrac{e^{\frac{1}{x}}}{x^2}$, ($\iota$) $\dfrac{x}{x^2 + a^2}$,

(κ) $x^3\sqrt{ax^2 + b}$, (λ) $\tan x \sec^2 x$, (μ) $\tan x \log \sin x$,

(ν) $\dfrac{(x^{-1} - 1)^5}{x^2}$, ($o$) $\dfrac{\tanh^{-1}x}{1 - x^2}$, ($\pi$) $\dfrac{2 + \log x}{x}$,

(ρ) $a^{1 + \sin x}\cos x$, (σ) $\dfrac{\sin x}{\sqrt{\cos x}}$, ($\tau$) $\dfrac{1}{\sqrt{1 - x^2}\sin^{-1}x}$.

4. Integrate after making appropriate changes such as $\sin^2 x = \frac{1}{2} - \frac{1}{2}\cos 2x$ or $\sec^2 x = 1 + \tan^2 x$, division of denominator into numerator, resolution of the product of trigonometric functions into a sum, completing the square, and so on.

(α) $\cos^2 2x$, (β) $\sin^4 x$. (γ) $\tan^4 x$,

(δ) $\dfrac{1}{x^2 + 3x + 25}$, (ϵ) $\dfrac{2x + 1}{x + 2}$, (ζ) $\dfrac{1 - \sin x}{\operatorname{vers} x}$,

(η) $\dfrac{x + 3}{4x^2 - 5x + 1}$, ($\theta$) $\dfrac{e^{2x} + e^x}{e^{2x} + 1}$, ($\iota$) $\dfrac{1}{\sqrt{2ax + x^2}}$,

(κ) $\sin 5x \cos 2x + 1$, (λ) $\sinh mx \sinh nx$, (μ) $\cos x \cos 2x \cos 3x$,

(ν) $\sec^6 x \tan x - \sqrt{2x}$, ($o$) $\dfrac{cx + d}{x^2 + ax + b}$, ($\pi$) $-\dfrac{x^{m-1}}{(ax^m + b)^p}$.

* The use of differentials (§ 35) is perhaps more familiar than the use of derivatives.

$$z(x) = \int \frac{dz}{dx}dx = \int \frac{dz}{dy}\frac{dy}{dx}dx = \int \frac{dz}{dy}dy = z[y(x)].$$

Then $$\int \frac{2}{x}\log x\, dx = \int 2\log x\, d\log x = (\log x)^2.$$

The use of this notation is left optional with the reader; it has some advantages and some disadvantages. The essential thing is to keep clearly in mind the fact that the problem is to be inspected with a view to detecting the function which will differentiate into the given integrand.

5. How are the following types integrated?

(α) $\sin^m x \cos^n x$, m or n odd, or m and n even,
(β) $\tan^n x$ or $\cot^n x$ when n is an integer,
(γ) $\sec^n x$ or $\csc^n x$ when n is even,
(δ) $\tan^m x \sec^n x$ or $\cot^m x \csc^n x$, n even.

6. Explain the alternative forms in (54)–(56) with all detail possible.

7. Find (α) the area under the parabola $y^2 = 4px$ from $x = 0$ to $x = a$; also (β) the corresponding volume of revolution. Find (γ) the total volume of an ellipsoid of revolution (see Ex. 9, p. 10).

8. Show that the area under $y = \sin mx \sin nx$ or $y = \cos mx \cos nx$ from $x = 0$ to $x = \pi$ is zero if m and n are unequal integers but $\frac{1}{2}\pi$ if they are equal.

9. Find the sectorial area of $r = a \tan\phi$ between the radii $\phi = 0$ and $\phi = \frac{1}{4}\pi$.

10. Find the area of the (α) lemniscate $r^2 = a^2 \cos 2\phi$ and (β) cardioid $r = 1 - \cos\phi$.

11. By Ex. 10, p. 10, find the volumes of these solids. Be careful to choose the parallel planes so that $A(x)$ may be found easily.

(α) The part cut off from a right circular cylinder by a plane through a diameter of one base and tangent to the other. *Ans.* $2/3\pi$ of the whole volume.

(β) How much is cut off from a right circular cylinder by a plane tangent to its lower base and inclined at an angle θ to the plane of the base?

(γ) A circle of radius $b < a$ is revolved, about a line in its plane at a distance a from its center, to generate a ring. The volume of the ring is $2\pi^2 ab^2$.

(δ) The axes of two equal cylinders of revolution of radius r intersect at right angles. The volume common to the cylinders is $16r^3/3$.

12. If the cross section of a solid is $A(x) = a_0x^3 + a_1x^2 + a_2x + a_3$, a cubic in x, the volume of the solid between two parallel planes is $\frac{1}{6}h(B + 4M + B')$ where h is the altitude and B and B' are the bases and M is the middle section.

13. Show that $\int \frac{1}{1+x^2} = \tan^{-1}\frac{x+c}{1-cx}$.

13. Aids to integration. The majority of cases of integration which arise in simple applications of calculus may be treated by the method of § 12. Of the remaining cases a large number cannot be integrated at all in terms of the functions which have been treated up to this point. Thus it is impossible to express $\int \frac{1}{\sqrt{(1-x^2)(1-a^2x^2)}}$ in terms of elementary functions. One of the chief reasons for introducing a variety of new functions in higher analysis is to have means for effecting the integrations called for by important applications. The discussion of this matter cannot be taken up here. The problem of integration from an elementary point of view calls for the tabulation of some devices which will accomplish the integration for a wide variety of integrands integrable in terms of elementary functions. The devices which will be treated are:

Integration by parts,	Resolution into partial fractions,
Various substitutions,	Reference to tables of integrals.

Integration by parts is an application of (61) when written as

$$\int uv' = uv - \int u'v. \tag{61'}$$

That is, it may happen that the integrand can be written as the product uv' of two factors, where v' is integrable and where $u'v$ is also integrable. Then uv' is integrable. For instance, $\log x$ is not integrated by the fundamental formulas; but

$$\int \log x = \int \log x \cdot 1 = x \log x - \int x/x = x \log x - x.$$

Here $\log x$ is taken as u and 1 as v', so that v is x, u' is $1/x$, and $u'v = 1$ is immediately integrable. This method applies to the inverse trigonometric and hyperbolic functions. Another example is

$$\int x \sin x = - x \cos x + \int \cos x = \sin x - x \cos x.$$

Here if $x = u$ and $\sin x = v'$, both v' and $u'v = - \cos x$ are integrable. If the choice $\sin x = u$ and $x = v'$ had been made, v' would have been integrable but $u'v = \frac{1}{2} x^2 \cos x$ would have been less simple to integrate than the original integrand. Hence in applying integration by parts it is necessary to *look ahead* far enough to see that both v' and $u'v$ are integrable, or at any rate that v' is integrable and the integral of $u'v$ is simpler than the original integral.*

Frequently integration by parts has to be applied several times in succession. Thus

$$\begin{aligned} \int x^2 e^x &= x^2 e^x - \int 2 x e^x && \text{if } u = x^2,\ v' = e^x, \\ &= x^2 e^x - 2\left[x e^x - \int e^x\right] && \text{if } u = x,\ v' = e^x, \\ &= x^2 e^x - 2 x e^x + 2 e^x. \end{aligned}$$

Sometimes it may be applied in such a way as to lead back to the given integral and thus afford an equation from which that integral can be obtained by solution. For example,

$$\begin{aligned} \int e^x \cos x &= e^x \cos x + \int e^x \sin x && \text{if } u = \cos x,\ v' = e^x, \\ &= e^x \cos x + \left[e^x \sin x - \int e^x \cos x\right] && \text{if } u = \sin x,\ v' = e^x, \\ &= e^x (\cos x + \sin x) - \int e^x \cos x. \end{aligned}$$

Hence $$\int e^x \cos x = \tfrac{1}{2} e^x (\cos x + \sin x).$$

* The method of differentials may again be introduced if desired.

14. For the *integration of a rational fraction* $f(x)/F(x)$ where f and F are polynomials in x, the fraction is first resolved into *partial fractions*. This is accomplished as follows. First if f is not of lower degree than F, divide F into f until the remainder is of lower degree than F. The fraction f/F is thus resolved into the sum of a polynomial (the quotient) and a fraction (the remainder divided by F) of which the numerator is of lower degree than the denominator. As the polynomial is integrable, it is merely necessary to consider fractions f/F where f is of lower degree than F. Next it is a fundamental theorem of algebra that a polynomial F may be resolved into linear and quadratic factors

$$F(x) = k(x-a)^\alpha (x-b)^\beta (x-c)^\gamma \cdots (x^2+mx+n)^\mu (x^2+px+q)^\nu \cdots,$$

where $a, b, c, \cdots$ are the real roots of the equation $F(x)=0$ and are of the respective multiplicities $\alpha, \beta, \gamma, \cdots$, and where the quadratic factors when set equal to zero give the pairs of conjugate imaginary roots of $F=0$, the multiplicities of the imaginary roots being $\mu, \nu, \cdots$. It is then a further theorem of algebra that the fraction f/F may be written as

$$\frac{f(x)}{F(x)} = \frac{A_1}{x-a} + \frac{A_2}{(x-a)^2} + \cdots + \frac{A_\alpha}{(x-a)^\alpha} + \frac{B_1}{x-b} + \cdots + \frac{B_\beta}{(x-b)^\beta} + \cdots$$
$$+ \frac{M_1x+N_1}{x^2+mx+n} + \frac{M_2x+N_2}{(x^2+mx+n)^2} + \cdots + \frac{M_\mu x+N_\mu}{(x^2+mx+n)^\mu} + \cdots,$$

where there is for each irreducible factor of F a term corresponding to the highest power to which that factor occurs in F and also a term corresponding to every lesser power. The coefficients $A, B, \cdots, M, N, \cdots$ may be obtained by clearing of fractions and equating coefficients of like powers of x, and solving the equations; or they may be obtained by clearing of fractions, substituting for x as many different values as the degree of F, and solving the resulting equations.

When f/F has thus been resolved into partial fractions, the problem has been reduced to the integration of each fraction, and this does not present serious difficulty. The following two examples will illustrate the method of resolution into partial fractions and of integration. Let it be required to integrate

$$\int \frac{x^2+1}{x(x-1)(x-2)(x^2+x+1)} \quad \text{and} \quad \int \frac{2x^3+6}{(x-1)^2(x-3)^3}.$$

The first fraction is expansible into partial fractions in the form

$$\frac{x^2+1}{x(x-1)(x-2)(x^2+x+1)} = \frac{A}{x} + \frac{B}{x-1} + \frac{C}{x-2} + \frac{Dx+E}{x^2+x+1}.$$

Hence $$x^2+1 = A(x-1)(x-2)(x^2+x+1) + Bx(x-2)(x^2+x+1)$$
$$+ Cx(x-1)(x^2+x+1) + (Dx+E)x(x-1)(x-2).$$

Rather than multiply out and equate coefficients, let $0, 1, 2, -1, -2$ be substituted. Then

$$1 = 2A, \quad 2 = -3B, \quad 5 = 14C, \quad D - E = 1/21, \quad E - 2D = 1/7,$$

$$\int \frac{x^2+1}{x(x-1)(x-2)(x^2+x+1)} = \int \frac{1}{2x} - \int \frac{2}{3(x-1)} + \int \frac{5}{14(x-2)} - \int \frac{4x+5}{21(x^2+x+1)}$$
$$= \frac{1}{2}\log x - \frac{2}{3}\log(x-1) + \frac{5}{14}\log(x-2) - \frac{2}{21}\log(x^2+x+1) - \frac{2}{7\sqrt{3}}\tan^{-1}\frac{2x+1}{\sqrt{3}}.$$

In the second case the form to be assumed for the expansion is

$$\frac{2x^3+6}{(x-1)^2(x-3)^3}=\frac{A}{x-1}+\frac{B}{(x-1)^2}+\frac{C}{(x-3)}+\frac{D}{(x-3)^2}+\frac{E}{(x-3)^3}.$$

$$2x^3+6=A(x-1)(x-3)^3+B(x-3)^3+C(x-1)^2(x-3)^2$$
$$+D(x-1)^2(x-3)+E(x-1)^2.$$

The substitution of 1, 3, 0, 2, 4 gives the equations

$$8=-8B,\qquad 60=4E,\qquad 9A+3C-D+12=0,$$
$$A-C+D+6=0,\qquad A+3C+3D=0.$$

The solutions are $-9/4$, -1, $+9/4$, $-3/2$, 15, and the integral becomes

$$\int\frac{2x^3+6}{(x-1)^2(x-3)^3}=-\frac{9}{4}\log(x-1)+\frac{1}{x-1}+\frac{9}{4}\log(x-3)$$
$$+\frac{3}{2(x-3)}-\frac{15}{2(x-3)^2}.$$

The importance of the fact that the method of partial fractions shows that *any rational fraction may be integrated* and, moreover, that the integral may at most consist of a rational part plus the logarithm of a rational fraction plus the inverse tangent of a rational fraction should not be overlooked. Taken with the method of substitution it establishes very wide categories of integrands which are integrable in terms of elementary functions, and effects their integration even though by a somewhat laborious method.

15. The *method of substitution* depends on the identity

$$\int_x f(x)=\int_y f[\phi(y)]\frac{dx}{dy}\qquad\text{if}\qquad x=\phi(y),\tag{59'}$$

which is allied to (59). To show that the integral on the right with respect to y is the integral of $f(x)$ with respect to x it is merely necessary to show that its derivative with respect to x is $f(x)$. By definition of integration,

$$\frac{d}{dy}\int_y f[\phi(y)]\frac{dx}{dy}=f[\phi(y)]\frac{dx}{dy}$$

and

$$\frac{d}{dx}\int_y f[\phi(y)]\frac{dx}{dy}=f[\phi(y)]\frac{dx}{dy}\cdot\frac{dy}{dx}=f[\phi(y)]$$

by (4). The identity is therefore proved. The method of integration by substitution is in fact seen to be merely such a systematization of the method based on (59) and set forth in § 12 as will make it practicable for more complicated problems. Again, differentials may be used if preferred.

Let R denote a rational function. To effect the integration of

$$\int\sin x\,R(\sin^2x,\cos x),\quad\text{let}\quad\cos x=y,\quad\text{then}\quad\int_y -R(1-y^2,y);$$

$$\int\cos x\,R(\cos^2x,\sin x),\quad\text{let}\quad\sin x=y,\quad\text{then}\quad\int_y R(1-y^2,y);$$

$$\int R\left(\frac{\sin x}{\cos x}\right)=\int R(\tan x),\quad\text{let}\quad\tan x=y,\quad\text{then}\quad\int_y\frac{R(y)}{1+y^2};$$

$$\int R(\sin x,\cos x),\quad\text{let}\quad\tan\frac{x}{2}=y,\quad\text{then}\quad\int_y R\left(\frac{2y}{1+y^2},\frac{1-y^2}{1+y^2}\right)\frac{2}{1+y^2}.$$

The last substitution renders any rational function of $\sin x$ and $\cos x$ rational in the variable y; it should not be used, however, if the previous ones are applicable — it is almost certain to give a more difficult final rational fraction to integrate.

A large number of geometric problems give integrands which are rational in x and in some one of the radicals $\sqrt{a^2+x^2}$, $\sqrt{a^2-x^2}$, $\sqrt{x^2-a^2}$. These may be converted into trigonometric or hyperbolic integrands by the following substitutions:

$$\int R(x, \sqrt{a^2-x^2}) \qquad x = a\sin y, \qquad \int_y R(a\sin y, a\cos y)\, a\cos y;$$

$$\int R(x, \sqrt{a^2+x^2}) \quad \begin{cases} x = a\tan y, & \int_y R(a\tan y, a\sec y)\, a\sec^2 y \\ x = a\sinh y, & \int_y R(a\sinh y, a\cosh y)\, a\cosh y; \end{cases}$$

$$\int R(x, \sqrt{x^2-a^2}) \quad \begin{cases} x = a\sec y, & \int_y R(a\sec y, a\tan y)\, a\sec y\tan y \\ x = a\cosh y, & \int_y R(a\cosh y, a\sinh y)\, a\sinh y. \end{cases}$$

It frequently turns out that the integrals on the right are easily obtained by methods already given; otherwise they can be treated by the substitutions above.

In addition to these substitutions there are a large number of others which are applied under specific conditions. Many of them will be found among the exercises. Moreover, it frequently happens that an integrand, which does not come under any of the standard types for which substitutions are indicated, is none the less integrable by some substitution which the form of the integrand will suggest.

Tables of integrals, giving the integrals of a large number of integrands, have been constructed by using various methods of integration. B. O. Peirce's "Short Table of Integrals" may be cited. If the particular integrand which is desired does not occur in the Table, it may be possible to devise some substitution which will reduce it to a tabulated form. In the Table are also given a large number of reduction formulas (for the most part deduced by means of integration by parts) which accomplish the successive simplification of integrands which could perhaps be treated by other methods, but only with an excessive amount of labor. Several of these reduction formulas are cited among the exercises. Although the Table is useful in performing integrations and indeed makes it to a large extent unnecessary to learn the various methods of integration, the exercises immediately below, which are constructed for the purpose of illustrating methods of integration, should be done without the aid of a Table.

EXERCISES

1. Integrate the following by parts:

$(\alpha)\ \int x\cosh x,\qquad (\beta)\ \int \tan^{-1}x,\qquad (\gamma)\ \int x^m\log x,$

$(\delta)\ \int \frac{\sin^{-1}x}{x^2},\qquad (\epsilon)\ \int \frac{xe^x}{(1+x)^2},\qquad (\zeta)\ \int \frac{1}{x(x^2-a^2)^{\frac{3}{2}}}.$

2. If $P(x)$ is a polynomial and $P'(x)$, $P''(x)$, $\cdots$ its derivatives, show

$$(\alpha)\ \int P(x)e^{ax} = \frac{1}{a}e^{ax}\left[P(x) - \frac{1}{a}P'(x) + \frac{1}{a^2}P''(x) - \cdots\right],$$

$$(\beta)\ \int P(x)\cos ax = \frac{1}{a}\sin ax\left[P(x) - \frac{1}{a^2}P''(x) + \frac{1}{a^4}P^{\text{iv}}(x) - \cdots\right]$$
$$+ \frac{1}{a}\cos ax\left[\frac{1}{a}P'(x) - \frac{1}{a^3}P'''(x) + \frac{1}{a^5}P^{\text{v}}(x) - \cdots\right],$$

and (γ) derive a similar result for the integrand $P(x)\sin ax$.

3. By successive integration by parts and subsequent solution, show

$(\alpha)\ \int e^{ax}\sin bx = \dfrac{e^{ax}(a\sin bx - b\cos bx)}{a^2+b^2}$,

$(\beta)\ \int e^{ax}\cos bx = \dfrac{e^{ax}(b\sin bx + a\cos bx)}{a^2+b^2}$,

$(\gamma)\ \int xe^{2x}\cos x = \frac{1}{25}e^{2x}[5x(\sin x + 2\cos x) - 4\sin x - 3\cos x]$.

4. Prove by integration by parts the reduction formulas

$(\alpha)\ \int \sin^m x\cos^n x = \dfrac{\sin^{m+1}x\cos^{n-1}x}{m+n} + \dfrac{n-1}{m+n}\int \sin^m x\cos^{n-2}x$,

$(\beta)\ \int \tan^m x\sec^n x = \dfrac{\tan^{m-1}x\sec^n x}{m+n-1} - \dfrac{m-1}{m+n-1}\int \tan^{n-2}x\sec^n x$,

$(\gamma)\ \int \dfrac{1}{(x^2+a^2)^n} = \dfrac{1}{2(n-1)a^2}\left[\dfrac{x}{(x^2+a^2)^{n-1}} + (2n-3)\int\dfrac{1}{(x^2+a^2)^{n-1}}\right]$,

$(\delta)\ \int \dfrac{x^m}{(\log x)^n} = -\dfrac{x^{m+1}}{(n-1)(\log x)^{n-1}} + \dfrac{m+1}{n-1}\int\dfrac{x^m}{(\log x)^{n-1}}$.

5. Integrate by decomposition into partial fractions:

$(\alpha)\ \int\dfrac{x^2-3x+3}{(x-1)(x-2)}$, $(\beta)\ \int\dfrac{1}{a^4-x^4}$, $(\gamma)\ \int\dfrac{1}{1+x^4}$,

$(\delta)\ \int\dfrac{x^2}{(x+2)^2(x+1)}$, $(\epsilon)\ \int\dfrac{4x^2-3x+1}{2x^5+x^3}$, $(\zeta)\ \int\dfrac{1}{x(1+x^2)^2}$.

6. Integrate by trigonometric or hyperbolic substitution:

$(\alpha)\ \int\sqrt{a^2-x^2}$, $(\beta)\ \int\sqrt{x^2-a^2}$, $(\gamma)\ \int\sqrt{a^2+x^2}$,

$(\delta)\ \int\dfrac{1}{(a-x^2)^{\frac{3}{2}}}$, $(\epsilon)\ \int\dfrac{\sqrt{x^2-a^2}}{x}$, $(\zeta)\ \int\dfrac{(a^{\frac{2}{3}}-x^{\frac{2}{3}})^{\frac{3}{2}}}{x^{\frac{1}{3}}}$.

7. Find the areas of these curves and their volumes of revolution:

$(\alpha)\ x^{\frac{2}{3}} + y^{\frac{2}{3}} = a^{\frac{2}{3}}$, $(\beta)\ a^4y^2 = a^2x^4 - x^6$, $(\gamma)\ \left(\dfrac{x}{a}\right)^2 + \left(\dfrac{y}{b}\right)^{\frac{2}{3}} = 1$.

8. Integrate by converting to a rational algebraic fraction:

$(\alpha)\ \int\dfrac{\sin 3x}{a^2\cos^2x+b^2\sin^2x}$, $(\beta)\ \int\dfrac{\cos 3x}{a^2\cos^2x+b^2\sin^2x}$, $(\gamma)\ \int\dfrac{\sin 2x}{a^2\cos^2x+b^2\sin^2x}$,

$(\delta)\ \int\dfrac{1}{a+b\cos x}$, $(\epsilon)\ \int\dfrac{1}{a+b\cos x+c\sin x}$, $(\zeta)\ \int\dfrac{1-\cos x}{1+\sin x}$.

9. Show that $\int R(x, \sqrt{a+bx+cx^2})$ may be treated by trigonometric substitution; distinguish between $b^2 - 4ac \gtrless 0$.

10. Show that $\int R\left(x, \sqrt[n]{\dfrac{ax+b}{cx+d}}\right)$ is made rational by $y^n = \dfrac{ax+b}{cx+d}$. Hence infer that $\int R(x, \sqrt{(x-\alpha)(x-\beta)})$ is rationalized by $y^2 = \dfrac{x-\beta}{x-\alpha}$. This accomplishes the integration of $R(x, \sqrt{a+bx+cx^2})$ when the roots of $a+bx+cx^2=0$ are real, that is, when $b^2-4ac>0$.

11. Show that $\int R\left[x, \left(\frac{ax+b}{cx+d}\right)^m, \left(\frac{ax+b}{cx+d}\right)^n, \cdots\right]$, where the exponents m, n, $\cdots$ are rational, is rationalized by $y^k = \frac{ax+b}{cx+d}$ if k is so chosen that km, kn, $\cdots$ are integers.

12. Show that $\int (a + by)^p y^q$ may be rationalized if p or q or $p + q$ is an integer. By setting $x^n = y$ show that $\int x^m (a + bx^n)^p$ may be reduced to the above type and hence is integrable when $\frac{m+1}{n}$ or p or $\frac{m+1}{n} + p$ is integral.

13. If the roots of $a + bx + cx^2 = 0$ are imaginary, $\int R(x, \sqrt{a + bx + cx^2})$ may be rationalized by $y = \sqrt{a + bx + cx^2} \mp x\sqrt{c}$.

14. Integrate the following.

$(\alpha)\ \int \frac{x^3}{\sqrt{x-1}}$, $\quad(\beta)\ \int \frac{1+\sqrt[3]{x}}{1+\sqrt[4]{x}}$, $\quad(\gamma)\ \int \frac{x}{\sqrt[3]{1+x}-\sqrt{1+x}}$,

$(\delta)\ \int \frac{e^{2x}}{\sqrt[4]{e^x+1}}$, $\quad(\epsilon)\ \int \frac{x^4}{\sqrt{(1-x^2)^3}}$, $\quad(\zeta)\ \int \frac{1}{(x-d)\sqrt{a+bx+cx^2}}$,

$(\eta)\ \int \frac{1}{x(1+x^2)^{\frac{3}{2}}}$, $\quad(\theta)\ \int \frac{\sqrt{2x^2+x}}{x^2}$, $\quad(\lambda)\ \int \frac{x^8}{\sqrt{1-x^3}} + \frac{\sqrt{1-x^3}}{x}$.

15. In view of Ex. 12 discuss the integrability of:

$(\alpha)\ \int \sin^m x \cos^n x$, let $\sin x = \sqrt{y}$, $\quad(\beta)\ \int \frac{x^m}{\sqrt{ax - x^2}}$ $\begin{cases}\text{let } x = ay^2,\\ \text{or } \sqrt{ax - x^2} = xy.\end{cases}$

16. Apply the reduction formulas, Table, p. 66, to show that the final integral for

$$\int \frac{x^m}{\sqrt{1-x^2}} \quad \text{is} \quad \int \frac{1}{\sqrt{1-x^2}} \quad \text{or} \quad \int \frac{x}{\sqrt{1-x^2}} \quad \text{or} \quad \int \frac{1}{x\sqrt{1-x^2}}$$

according as m is even or odd and positive or odd and negative.

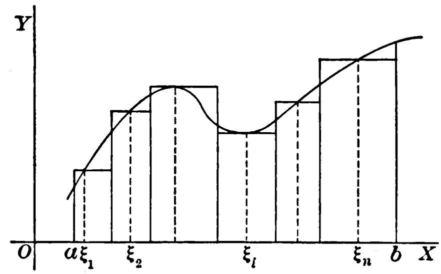
17. Prove sundry of the formulas of Peirce's Table.

18. Show that if $R(x, \sqrt{a^2 - x^2})$ contains x only to odd powers, the substitution $z = \sqrt{a^2 - x^2}$ will rationalize the expression. Use Exs. 1 (ζ) and 6 (ϵ) to compare the labor of this algebraic substitution with that of the trigonometric or hyperbolic.

16. Definite integrals. If an interval from $x = a$ to $x = b$ be divided into n successive intervals Δx_1, Δx_2, $\cdots$, Δx_n and the value $f(\xi_i)$ of a function $f(x)$ be computed from some point ξ_i in each interval Δx_i and be multiplied by Δx_i, then *the limit of the sum*

$$\lim_{\substack{\Delta x_i \doteq 0 \\ n=\infty}} [f(\xi_1)\Delta x_1 + f(\xi_2)\Delta x_2 + \cdots + f(\xi_n)\Delta x_n] = \int_a^b f(x)\,dx. \qquad (62)$$

when each interval becomes infinitely short and their number n becomes infinite, is known as *the definite integral* of $f(x)$ from a to b, and is designated as indicated. If $y=f(x)$ be graphed, the sum will be represented by the area under a broken line, and it is clear that the limit of the sum, that is, the integral, will be represented by the *area under the curve* $y=f(x)$ and between the ordinates $x=a$ and $x=b$. Thus the definite integral, defined arithmetically by (62), may be connected with a geometric concept which can serve to suggest properties of the integral much as the interpretation of the derivative as the slope of the tangent served as a useful geometric representation of the arithmetical definition (2).



For instance, if a, b, c are successive values of x, then

$$\int_a^b f(x)\,dx + \int_b^c f(x)\,dx = \int_a^c f(x)\,dx \tag{63}$$

is the equivalent of the fact that the area from a to c is equal to the sum of the areas from a to b and b to c. Again, if Δx be considered positive when x moves from a to b, it must be considered negative when x moves from b to a and hence from (62)

$$\int_a^b f(x)\,dx = -\int_b^a f(x)\,dx. \tag{64}$$

Finally, if M be the maximum of $f(x)$ in the interval, the area under the curve will be less than that under the line $y=M$ through the highest point of the curve; and if m be the minimum of $f(x)$, the area under the curve is greater than that under $y=m$. Hence

$$m(b-a) < \int_a^b f(x)\,dx < M(b-a). \tag{65}$$

There is, then, some intermediate value $m<\mu<M$ such that the integral is equal to $\mu(b-a)$; and if the line $y=\mu$ cuts the curve in a point whose abscissa is ξ intermediate between a and b, then

$$\int_a^b f(x)\,dx = \mu(b-a) = (b-a)f(\xi). \tag{65'}$$

This is the fundamental *Theorem of the Mean* for definite integrals.

The definition (62) may be applied directly to the evaluation of the definite integrals of the simplest functions. Consider first $1/x$ and let a, b be positive with a less than b. Let the interval from a to b be divided into n intervals Δx_i which are in geometrical progression in the ratio r so that $x_1 = a$, $x_2 = ar$, $\cdots$, $x_{n+1} = ar^n$

and $\Delta x_1 = a(r-1)$, $\Delta x_2 = ar(r-1)$, $\Delta x_3 = ar^2(r-1)$, $\cdots$, $\Delta x_n = ar^{n-1}(r-1)$;

whence $\quad b - a = \Delta x_1 + \Delta x_2 + \cdots + \Delta x_n = a(r^n - 1)$ and $r^n = b/a$.

Choose the points ξ_i in the intervals Δx_i as the initial points of the intervals. Then

$$\frac{\Delta x_1}{\xi_1} + \frac{\Delta x_2}{\xi_2} + \cdots + \frac{\Delta x_n}{\xi_n} = \frac{a(r-1)}{a} + \frac{ar(r-1)}{ar} + \cdots + \frac{ar^{n-1}(r-1)}{ar^{n-1}} = n(r-1).$$

But $\qquad r = \sqrt[n]{b/a} \quad$ or $\quad n = \log(b/a) \div \log r$.

Hence $\quad \frac{\Delta x_1}{\xi_1} + \frac{\Delta x_2}{\xi_2} + \cdots + \frac{\Delta x_n}{\xi_n} = n(r-1) = \log\frac{b}{a}\cdot\frac{r-1}{\log r} = \log\frac{b}{a}\cdot\frac{h}{\log(1+h)}.$

Now if n becomes infinite, r approaches 1, and h approaches 0. But the limit of $\log(1+h)/h$ as $h \doteq 0$ is by definition the derivative of $\log(1+x)$ when $x = 0$ and is 1. Hence

$$\int_a^b \frac{dx}{x} = \lim_{n=\infty}\left[\frac{\Delta x_1}{\xi_1} + \frac{\Delta x_2}{\xi_2} + \cdots + \frac{\Delta x_n}{\xi_n}\right] = \log\frac{b}{a} = \log b - \log a.$$

As another illustration let it be required to evaluate the integral of $\cos^2 x$ from 0 to $\frac{1}{2}\pi$. Here let the intervals Δx_i be equal and their number odd. Choose the ξ's as the initial points of their intervals. The sum of which the limit is desired is

$$\sigma = \cos^2 0\cdot\Delta x + \cos^2\Delta x\cdot\Delta x + \cos^2 2\,\Delta x\cdot\Delta x + \cdots \\ + \cos^2(n-2)\,\Delta x\cdot\Delta x + \cos^2(n-1)\,\Delta x\cdot\Delta x.$$

But $\quad n\Delta x = \frac{1}{2}\pi$, and $(n-1)\,\Delta x = \frac{1}{2}\pi - \Delta x$, $(n-2)\,\Delta x = \frac{1}{2}\pi - 2\,\Delta x$, $\cdots$,

and $\qquad \cos(\frac{1}{2}\pi - y) = \sin y \quad$ and $\quad \sin^2 y + \cos^2 y = 1$.

Hence $\quad \sigma = \Delta x\,[\cos^2 0 + \cos^2\Delta x + \cos^2 2\,\Delta x + \cdots + \sin^2 2\,\Delta x + \sin^2\Delta x]$
$$= \Delta x\left[1 + \frac{n-1}{2}\right].$$

Hence $\quad \int_0^{\frac{\pi}{2}} \cos^2 x\,dx = \lim_{\Delta x \doteq 0}\left[\frac{1}{2}n\,\Delta x + \frac{1}{2}\,\Delta x\right] = \lim_{\Delta x \doteq 0}\left(\frac{1}{4}\pi + \frac{1}{2}\,\Delta x\right) = \frac{1}{4}\pi.$

Indications for finding the integrals of other functions are given in the exercises.

It should be noticed that the variable x which appears in the expression of the definite integral really has nothing to do with the value of the integral but merely serves as a symbol useful in forming the sum in (62). What is of importance is the function f and the limits a, b of the interval over which the integral is taken.

$$\int_a^b f(x)\,dx = \int_a^b f(t)\,dt = \int_a^b f(y)\,dy = \int_a^b f(*)\,d*.$$

The variable in the integrand disappears in the integration and leaves the value of the integral as a function of the limits a and b alone.

17. If the lower limit of the integral be fixed, the value

$$\int_a^b f(x)\,dx = \Phi(b)$$

of the integral is a function of the upper limit regarded as variable. To find the derivative $\Phi'(b)$, form the quotient (2),

$$\frac{\Phi(b+\Delta b)-\Phi(b)}{\Delta b} = \frac{\int_a^{b+\Delta b} f(x)\,dx - \int_a^b f(x)\,dx}{\Delta b}.$$

By applying (63) and (65′), this takes the simpler form

$$\frac{\Phi(b+\Delta b)-\Phi(b)}{\Delta b} = \frac{\int_b^{b+\Delta b} f(x)\,dx}{\Delta b} = \frac{1}{\Delta b}\cdot f(\xi)\,\Delta b,$$

where ξ is intermediate between b and $b+\Delta b$. Let $\Delta b \doteq 0$. Then ξ approaches a and $f(\xi)$ approaches $f(a)$. Hence

$$\Phi'(b) = \frac{d}{db}\int_a^b f(x)\,dx = f(b). \qquad (66)$$

If preferred, the variable b may be written as x, and

$$\Phi(x) = \int_a^x f(x)\,dx, \qquad \Phi'(x) = \frac{d}{dx}\int_a^x f(x)\,dx = f(x). \qquad (66')$$

This equation will establish the relation between the definite integral and the indefinite integral. For by definition, the indefinite integral $F(x)$ of $f(x)$ is any function such that $F'(x)$ equals $f(x)$. As $\Phi'(x) = f(x)$ it follows that

$$\int_a^x f(x)\,dx = F(x) + C. \qquad (67)$$

Hence except for an additive constant, the indefinite integral of f is the definite integral of f from a fixed lower limit to a variable upper limit. As the definite integral vanishes when the upper limit coincides with the lower, the constant C is $-F(a)$ and

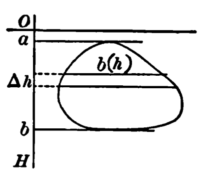
$$\int_a^b f(x)\,dx = F(b) - F(a). \qquad (67')$$

Hence, *the definite integral of $f(x)$ from a to b is the difference between the values of any indefinite integral $F(x)$ taken for the upper and lower limits of the definite integral;* and if the indefinite integral of f is known, the definite integral may be obtained without applying the definition (62) to f.

The great importance of definite integrals to geometry and physics lies in that fact that *many quantities* connected with geometric figures or physical bodies *may be expressed simply for small portions* of the figures or bodies and may then be obtained as the sum of those quantities taken over all the small portions, or rather, as the *limit of that sum when the portions become smaller and smaller.* Thus the area under a curve cannot in the first instance be evaluated; but if only that portion of the curve which lies over a small interval Δx be considered and the rectangle corresponding to the ordinate $f(\xi)$ be drawn, it is clear that the area of the rectangle is $f(\xi)\,\Delta x$, that the area of all the rectangles is the sum $\Sigma f(\xi)\,\Delta x$ taken from a to b, that when the intervals Δx approach zero the limit of their sum is the area under the curve; and hence that area may be written as the definite integral of $f(x)$ from a to b.*

In like manner consider *the mass of a rod* of variable density and suppose the rod to lie along the x-axis so that the density may be taken as a function of x. In any small length Δx of the rod the density is nearly constant and the mass of that part is approximately equal to the product $\rho\Delta x$ of the density $\rho(x)$ at the initial point of that part times the length Δx of the part. In fact it is clear that the mass will be intermediate between the products $m\Delta x$ and $M\Delta x$, where m and M are the minimum and maximum densities in the interval Δx. In other words the mass of the section Δx will be exactly equal to $\rho(\xi)\,\Delta x$ where ξ is some value of x in the interval Δx. The mass of the whole rod is therefore the sum $\Sigma\rho(\xi)\Delta x$ taken from one end of the rod to the other, and if the intervals be allowed to approach zero, the mass may be written as the integral of $\rho(x)$ from one end of the rod to the other.†

Another problem that may be treated by these methods is that of finding the *total pressure* on a vertical area submerged in a liquid, say, in water. Let w be the weight of a column of water of cross section 1 sq. unit and of height 1 unit. (If the unit is a foot, $w = 62.5$ lb.) At a point h units below the surface of the water the pressure is wh and upon a small area near that depth the pressure is approximately whA if A be the area. The pressure on the area A is exactly equal to $w\xi A$ if ξ is some depth intermediate between that of the top and that of the bottom of the area. Now let the finite area be ruled into strips of height Δh. Consider the product $wh\,b(h)\,\Delta h$ where $b(h) = f(h)$ is the breadth of the area at the depth h. This



* The ξ's may evidently be so chosen that the finite sum $\Sigma f(\xi)\,\Delta x$ is exactly equal to the area under the curve; but still it is necessary to let the intervals approach zero and thus replace the sum by an integral because the values of ξ which make the sum equal to the area are unknown.

† This and similar problems, here treated by using the Theorem of the Mean for integrals, may be treated from the point of view of differentiation as in § 7 or from that of Duhamel's or Osgood's Theorem as in §§ 34, 35. It should be needless to state that in any particular problem some one of the three methods is likely to be somewhat preferable to either of the others. The reason for laying such emphasis upon the Theorem of the Mean here and in the exercises below is that the theorem is in itself very important and needs to be thoroughly mastered.

is approximately the pressure on the strip as it is the pressure at the top of the strip multiplied by the approximate area of the strip. Then $w\xi b(\xi)\Delta h$, where ξ is some value between h and $h + \Delta h$, is the actual pressure on the strip. (It is sufficient to write the pressure as approximately $whb(h)\Delta h$ and not trouble with the ξ.) The total pressure is then $\Sigma w\xi b(\xi)\Delta h$ or better the limit of that sum. Then

$$P = \lim \sum w\xi b(\xi)\, dh = \int_a^b whb(h)\, dh,$$

where a is the depth of the top of the area and b that of the bottom. To evaluate the pressure it is merely necessary to find the breadth b as a function of h and integrate.

EXERCISES

1. If k is a constant, show $\int_a^b kf(x)\,dx = k\int_a^b f(x)\,dx$.

2. Show that $\int_a^b (u \pm v)\,dx = \int_a^b u dx \pm \int_a^b v dx$.

3. If, from a to b, $\psi(x) < f(x) < \phi(x)$, show $\int_a^b \psi(x)\,dx < \int_a^b f(x)\,dx < \int_a^b \phi(x)\,dx$.

4. Suppose that the minimum and maximum of the quotient $Q(x) = f(x)/\phi(x)$ of two functions in the interval from a to b are m and M, and let $\phi(x)$ be positive so that

$$m < Q(x) = \frac{f(x)}{\phi(x)} < M \quad \text{and} \quad m\phi(x) < f(x) < M\phi(x)$$

are true relations. Show by Exs. 3 and 1 that

$$m < \frac{\int_a^b f(x)\,dx}{\int_a^b \phi(x)\,dx} < M \quad \text{and} \quad \frac{\int_a^b f(x)\,dx}{\int_a^b \phi(x)\,dx} = \mu = Q(\xi) = \frac{f(\xi)}{\phi(\xi)},$$

where ξ is some value of x between a and b.

5. If m and M are the minimum and maximum of $f(x)$ between a and b and if $\phi(x)$ is always positive in the interval, show that

$$m\int_a^b \phi(x)\,dx < \int_a^b f(x)\,\phi(x)\,dx < M\int_a^b \phi(x)\,dx$$

and

$$\int_a^b f(x)\,\phi(x)\,dx = \mu\int_a^b \phi(x)\,dx = f(\xi)\int_a^b \phi(x)\,dx.$$

Note that the integrals of $[M - f(x)]\,\phi(x)$ and $[f(x) - m]\,\phi(x)$ are positive and apply Ex. 2.

6. Evaluate the following by the direct application of (62):

$$(\alpha)\quad \int_a^b x dx = \frac{b^2 - a^2}{2}, \qquad (\beta)\quad \int_a^b e^x dx = e^b - e^a.$$

Take equal intervals and use the rules for arithmetic and geometric progressions.

7. Evaluate (α) $\int_a^b x^m dx = \frac{1}{m+1}(b^{m+1} - a^{m+1})$, $\quad(\beta)$ $\int_a^b c^x dx = \frac{1}{\log c}(c^b - c^a)$.
In the first the intervals should be taken in geometric progression with $r^n = b/a$.

8. Show directly that (α) $\int_0^{\pi} \sin^2 x dx = \frac{1}{2}\pi$, ($\beta$) $\int_0^{\pi} \cos^n x dx = 0$, if n is odd.

9. With the aid of the trigonometric formulas

$$\cos x + \cos 2x + \cdots + \cos(n-1)x = \tfrac{1}{2}[\sin nx \cot \tfrac{1}{2}x - 1 - \cos nx],$$
$$\sin x + \sin 2x + \cdots + \sin(n-1)x = \tfrac{1}{2}[(1 - \cos nx)\cot \tfrac{1}{2}x - \sin nx],$$

show (α) $\int_a^b \cos x dx = \sin b - \sin a$, ($\beta$) $\int_a^b \sin x dx = \cos a - \cos b$.

10. A function is said to be *even* if $f(-x) = f(x)$ and *odd* if $f(-x) = -f(x)$. Show (α) $\int_{-a}^{+a} f(x)\,dx = 2\int_0^a f(x)\,dx$, f even, (β) $\int_{-a}^{+a} f(x)\,dx = 0$, f odd.

11. Show that if an integral is regarded as a function of the lower limit, the upper limit being fixed, then

$$\Phi'(a) = \frac{d}{da}\int_a^b f(x)\,dx = -f(a), \quad \text{if } \Phi(a) = \int_a^b f(x)\,dx.$$

12. Use the relation between definite and indefinite integrals to compare

$$\int_a^b f(x)\,dx = (b-a)f(\xi) \quad \text{and} \quad F(b) - F(a) = (b-a)F'(\xi),$$

the Theorem of the Mean for derivatives and for definite integrals.

13. From consideration of Exs. 12 and 4 establish *Cauchy's Formula*

$$\frac{\Delta F}{\Delta \Phi} = \frac{F(b) - F(a)}{\Phi(b) - \Phi(a)} = \frac{F'(\xi)}{\Phi'(\xi)}, \qquad a < \xi < b,$$

which states that the quotient of the increments ΔF and $\Delta\Phi$ of two functions, in any interval in which the derivative $\Phi'(x)$ does not vanish, is equal to the quotient of the derivatives of the functions for some interior point of the interval. What would the application of the Theorem of the Mean for derivatives to numerator and denominator of the left-hand fraction give, and wherein does it differ from Cauchy's Formula?

14. Discuss the volume of revolution of $y = f(x)$ as the limit of the sum of thin cylinders and compare the results with those found in Ex. 9, p. 10.

15. Show that the mass of a rod running from a to b along the x-axis is $\frac{1}{2}k(b^2 - a^2)$ if the density varies as the distance from the origin (k is a factor of proportionality).

16. Show (α) that the mass in a rod running from a to b is the same as the area under the curve $y = \rho(x)$ between the ordinates $x = a$ and $x = b$, and explain why this should be seen intuitively to be so. Show (β) that if the density in a plane slab bounded by the x-axis, the curve $y = f(x)$, and the ordinates $x = a$ and $x = b$ is a function $\rho(x)$ of x alone, the mass of the slab is $\int_a^b y\rho(x)\,dx$; also (γ) that the mass of the corresponding volume of revolution is $\int_a^b \pi y^2 \rho(x)\,dx$.

17. An isosceles triangle has the altitude a and the base $2b$. Find (α) the mass on the assumption that the density varies as the distance from the vertex (measured along the altitude). Find (β) the mass of the cone of revolution formed by revolving the triangle about its altitude if the law of density is the same.

18. In a plane, the *moment of inertia I of a particle of mass m with respect to a point* is defined as the product mr^2 of the mass by the square of its distance from the point. Extend this definition from particles to bodies.

(α) Show that the moments of inertia of a rod running from a to b and of a circular slab of radius a are respectively

$$I = \int_a^b x^2\rho(x)\,dx \quad \text{and} \quad I = \int_0^a 2\pi r^3\rho(r)\,dr, \qquad \rho \text{ the density,}$$

if the point of reference for the rod is the origin and for the slab is the center.

(β) Show that for a rod of length $2l$ and of uniform density, $I = \frac{1}{3}Ml^2$ with respect to the center and $I = \frac{4}{3}Ml^2$ with respect to the end, M being the total mass of the rod.

(γ) For a uniform circular slab with respect to the center $I = \frac{1}{2}Ma^2$.

(δ) For a uniform rod of length $2l$ with respect to a point at a distance d from its center is $I = M(\frac{1}{3}l^2 + d^2)$. Take the rod along the axis and let the point be (α, β) with $d^2 = \alpha^2 + \beta^2$.

19. A rectangular gate holds in check the water in a reservoir. If the gate is submerged over a vertical distance H and has a breadth B and the top of the gate is a units below the surface of the water, find the pressure on the gate. At what depth in the water is the point where the pressure is the mean pressure over the gate?

20. A dam is in the form of an isosceles trapezoid 100 ft. along the top (which is at the water level) and 60 ft. along the bottom and 30 ft. high. Find the pressure in tons.

21. Find the pressure on a circular gate in a water main if the radius of the circle is r and the depth of the center of the circle below the water level is $d \geqq r$.

22. In space, *moments of inertia* are defined *relative to an axis* and in the formula $I = mr^2$, for a single particle, r is the perpendicular distance from the particle to the axis.

(α) Show that if the density in a solid of revolution generated by $y = f(x)$ varies only with the distance along the axis, the moment of inertia about the axis of revolution is $I = \int_a^b \frac{1}{2}\pi y^4\rho(x)\,dx$. Apply Ex. 18 after dividing the solid into disks.

(β) Find the moment of inertia of a sphere about a diameter in case the density is constant; $I = \frac{2}{5}Ma^2 = \frac{8}{15}\pi\rho a^5$.

(γ) Apply the result to find the moment of inertia of a spherical shell with external and internal radii a and b; $I = \frac{2}{5}M(a^5 - b^5)/(a^3 - b^3)$. Let $b \doteq a$ and thus find $I = \frac{2}{3}Ma^2$ as the moment of inertia of a spherical surface (shell of negligible thickness).

(δ) For a cone of revolution $I = \frac{3}{10}Ma^2$ where a is the radius of the base.

23. If the force of attraction exerted by a mass m upon a point is $kmf(r)$ where r is the distance from the mass to the point, show that the attraction exerted at the origin by a rod of density $\rho(x)$ running from a to b along the x-axis is

$$A = \int_a^b kf(x)\,\rho(x)\,dx, \quad \text{and that} \quad A = kM/ab, \qquad M = \rho(b - a),$$

is the attraction of a uniform rod if the law is the Law of Nature, that is, $f(r) = 1/r^2$.

24. Suppose that the density ρ in the slab of Ex. 16 were a function $\rho(x, y)$ of both x and y. Show that the mass of a small slice over the interval Δx_i would be of the form

$$\Delta x \int_0^{y=f(\xi)} \rho(x, y)\,dy = \Phi(\xi)\,\Delta x \quad \text{and that} \quad \int_a^b \Phi(x)\,\Delta x = \int_a^b \left[\int_0^{y=f(x)} \rho(x, y)\,dy\right] dx$$

would be the expression for the total mass and would require an integration with respect to y in which x was held constant, a substitution of the limits $f(x)$ and 0 for y, and then an integration with respect to x from a to b.

25. Apply the considerations of Ex. 24 to finding moments of inertia of
(α) a uniform triangle $y = mx$, $y = 0$, $x = a$ with respect to the origin,
(β) a uniform rectangle with respect to the center,
(γ) a uniform ellipse with respect to the center.

26. Compare Exs. 24 and 16 to treat the volume under the surface $z = \rho(x, y)$ and over the area bounded by $y = f(x)$, $y = 0$, $x = a$, $x = b$. Find the volume
(α) under $z = xy$ and over $y^2 = 4px$, $y = 0$, $x = 0$, $x = b$,
(β) under $z = x^2 + y^2$ and over $x^2 + y^2 = a^2$, $y = 0$, $x = 0$, $x = Q$,
(γ) under $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} + \dfrac{z^2}{c^2} = 1$ and over $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 1$, $y = 0$, $x = 0$, $x = a$.

27. Discuss sectorial area $\frac{1}{2}\int r^2 d\phi$ in polar coördinates as the limit of the sum of small sectors running out from the pole.

28. Show that the moment of inertia of a uniform circular sector of angle α and radius a is $\frac{1}{4}\rho\alpha a^4$. Hence infer $I = \frac{1}{4}\rho\int_{\alpha_0}^{\alpha_1} r^4 d\phi$ in polar coördinates.

29. Find the moment of inertia of a uniform (α) lemniscate $r^2 = a^2 \cos^2 2\phi$ and (β) cardioid $r = a(1 - \cos\phi)$ with respect to the pole. Also of (γ) the circle $r = 2a\cos\phi$ and (δ) the rose $r = a \sin 2\phi$ and (ϵ) the rose $r = a \sin 3\phi$.

CHAPTER II

REVIEW OF FUNDAMENTAL THEORY*

18. Numbers and limits. The concept and theory of *real number*, integral, rational, and irrational, will not be set forth in detail here. Some matters, however, which are necessary to the proper understanding of rigorous methods in analysis must be mentioned; and numerous points of view which are adopted in the study of irrational number will be suggested in the text or exercises.

It is taken for granted that by his earlier work the reader has become familiar with the use of real numbers. In particular it is assumed that he is accustomed to represent numbers as a *scale*, that is, by points on a straight line, and that he knows that when a line is given and an origin chosen upon it and a unit of measure and a positive direction have been chosen, then to each point of the line corresponds one and only one real number, and conversely. Owing to this correspondence, that is, owing to the conception of a scale, it is possible to interchange statements about numbers with statements about points and hence to obtain a more vivid and graphic or a more abstract and arithmetic phraseology as may be desired. Thus instead of saying that the numbers $x_1, x_2, \cdots$ are increasing algebraically, one may say that the points (whose coördinates are) $x_1, x_2, \cdots$ are moving in the positive direction or to the right; with a similar correlation of a decreasing suite of numbers with points moving in the negative direction or to the left. It should be remembered, however, that whether a statement is couched in geometric or algebraic terms, it is always a statement concerning numbers when one has in mind the point of view of pure analysis.†

It may be recalled that arithmetic begins with the integers, including 0, and with addition and multiplication. That second, the rational numbers of the form p/q are introduced with the operation of division and the negative rational numbers with the operation of subtraction. Finally, the irrational numbers are introduced by various processes. Thus $\sqrt{2}$ occurs in geometry through the necessity of expressing the length of the diagonal of a square, and $\sqrt{3}$ for the diagonal of a cube. Again, π is needed for the ratio of circumference to diameter in a circle. In algebra any equation of odd degree has at least one real root and hence may be regarded as defining a number. But there is an essential difference between rational and irrational numbers in that any rational number is of the

* The object of this chapter is to set forth systematically, with attention to precision of statement and accuracy of proof, those fundamental definitions and theorems which lie at the basis of calculus and which have been given in the previous chapter from an intuitive rather than a critical point of view.

† Some illustrative graphs will be given; the student should make many others.

form $\pm p/q$ with $q \neq 0$ and can therefore be written down explicitly; whereas the irrational numbers arise by a variety of processes and, although they may be represented to any desired accuracy by a decimal, they cannot all be written down explicitly. It is therefore necessary to have some definite axioms regulating the essential properties of irrational numbers. The particular axiom upon which stress will here be laid is the axiom of continuity, the use of which is essential to the proof of elementary theorems on limits.

19. AXIOM OF CONTINUITY. *If all the points of a line are divided into two classes such that every point of the first class precedes every point of the second class, there must be a point C such that any point preceding C is in the first class and any point succeeding C is in the second class.* This principle may be stated in terms of numbers, as: *If all real numbers be assorted into two classes such that every number of the first class is algebraically less than every number of the second class, there must be a number N such that any number less than N is in the first class and any number greater than N is in the second.* The number N (or point C) is called the frontier number (or point), or simply the *frontier* of the two classes, and in particular it is the *upper frontier* for the first class and the *lower frontier* for the second.

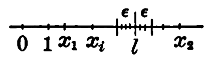
To consider a particular case, let all the negative numbers and zero constitute the first class and all the positive numbers the second, or let the negative numbers alone be the first class and the positive numbers with zero the second. In either case it is clear that the classes satisfy the conditions of the axiom and that zero is the frontier number such that any lesser number is in the first class and any greater in the second. If, however, one were to consider the system of all positive and negative numbers but without zero, it is clear that there would be no number N which would satisfy the conditions demanded by the axiom when the two classes were the negative and positive numbers; for no matter how small a positive number were taken as N, there would be smaller numbers which would also be positive and would not belong to the first class; and similarly in case it were attempted to find a negative N. Thus the axiom insures the presence of zero in the system, and in like manner insures the presence of every other number — a matter which is of importance because there is no way of writing all (irrational) numbers in explicit form.

Further to appreciate the continuity of the number scale, consider the four significations attributable to the phrase "*the interval from a to b.*" They are

$$a \leqq x \leqq b, \qquad a < x \leqq b, \qquad a \leqq x < b, \qquad a < x < b.$$

That is to say, both end points or either or neither may belong to the interval. In the case a is absent, the interval has no first point; and if b is absent, there is no last point. Thus if zero is not counted as a positive number, there is no least positive number; for if any least number were named, half of it would surely be less, and hence the absurdity. The axiom of continuity shows that if all numbers be divided into two classes as required, there must be either a greatest in the first class or a least in the second — the frontier — but not both unless the frontier is counted twice, once in each class.

20. Definition of a Limit. *If x is a variable which takes on successive values $x_1, x_2, \cdots, x_i, x_j, \cdots$, the variable x is said to approach the constant l as a limit if the numerical difference between x and l ultimately becomes, and for all succeeding values of x remains, less than any preassigned number no matter how small.* The numerical difference between x and l is denoted by $|x-l|$ or $|l-x|$ and is called the *absolute value* of the difference. The fact of the approach to a limit may be stated as



$$|x-l|<\epsilon \quad \text{for all } x\text{'s subsequent to some } x$$

or
$$x=l+\eta, \quad |\eta|<\epsilon \quad \text{for all } x\text{'s subsequent to some } x,$$

where ϵ is a positive number which may be assigned at pleasure and must be assigned before the attempt be made to find an x such that for all subsequent x's the relation $|x-l|<\epsilon$ holds.

So long as the conditions required in the definition of a limit are satisfied there is no need of bothering about how the variable approaches its limit, whether from one side or alternately from one side and the other, whether discontinuously as in the case of the area of the polygons used for computing the area of a circle or continuously as in the case of a train brought to rest by its brakes. To speak geometrically, a point x which changes its position upon a line approaches the point l as a limit if the point x ultimately comes into and remains in an assigned interval, no matter how small, surrounding l.

A variable is said to *become infinite* if the numerical value of the variable ultimately becomes and remains greater than any preassigned number K, no matter how large.* The notation is $x=\infty$, but had best be read "x becomes infinite," not "x equals infinity."

Theorem 1. If a variable is always increasing, it either becomes infinite or approaches a limit.

That the variable *may* increase indefinitely is apparent. But if it does not become infinite, there must be numbers K which are greater than any value of the variable. Then any number must satisfy one of two conditions: either there are values of the variable which are greater than it or there are no values of the variable greater than it. Moreover all numbers that satisfy the first condition are less than any number which satisfies the second. All numbers are therefore divided into two classes fulfilling the requirements of the axiom of continuity, and there must be a number N such that there are values of the variable greater than any number $N-\epsilon$ which is less than N. Hence if ϵ be assigned, there is a value of the variable which lies in the interval $N-\epsilon<x\leqq N$, and as the variable is always increasing, all subsequent values must lie in this interval. Therefore the variable approaches N as a limit.

* This definition means what it says, and no more. Later, additional or different meanings may be assigned to infinity, but not now. Loose and extraneous concepts in this connection are almost certain to introduce errors and confusion.

EXERCISES

1. If $x_1, x_2, \cdots, x_n, \cdots, x_{n+p}, \cdots$ is a suite approaching a limit, apply the definition of a limit to show that when ϵ is given it must be possible to find a value of n so great that $|x_{n+p} - x_n| < \epsilon$ for all values of p.

2. If $x_1, x_2, \cdots$ is a suite approaching a limit and if $y_1, y_2, \cdots$ is any suite such that $|y_n - x_n|$ approaches zero when n becomes infinite, show that the y's approach a limit which is identical with the limit of the x's.

3. As the definition of a limit is phrased in terms of inequalities and absolute values, note the following rules of operation:

$$(\alpha)\ \text{If}\quad a > 0 \quad \text{and} \quad c > b, \quad \text{then} \quad \frac{c}{a} > \frac{b}{a} \quad \text{and} \quad \frac{a}{c} < \frac{a}{b},$$

$$(\beta)\ |a + b + c + \cdots| \leqq |a| + |b| + |c| + \cdots, \qquad (\gamma)\ |abc \cdots| = |a| \cdot |b| \cdot |c| \cdots,$$

where the equality sign in (β) holds only if the numbers $a, b, c, \cdots$ have the same sign. By these relations and the definition of a limit prove the fundamental theorems:

If $\lim x = X$ and $\lim y = Y$, then $\lim (x \pm y) = X \pm Y$ and $\lim xy = XY$.

4. Prove Theorem 1 when restated in the slightly changed form: If a variable x never decreases and never exceeds K, then x approaches a limit N and $N \leqq K$. Illustrate fully. State and prove the corresponding theorem for the case of a variable never increasing.

5. If $x_1, x_2, \cdots$ and $y_1, y_2, \cdots$ are two suites of which the first never decreases and the second never increases, all the y's being greater than any of the x's, and if when ϵ is assigned an n can be found such that $y_n - x_n < \epsilon$, show that the limits of the suites are identical.

6. If $x_1, x_2, \cdots$ and $y_1, y_2, \cdots$ are two suites which never decrease, show by Ex. 4 (not by Ex. 3) that the suites $x_1 + y_1, x_2 + y_2, \cdots$ and $x_1y_1, x_2y_2, \cdots$ approach limits. Note that two infinite decimals are precisely two suites which never decrease as more and more figures are taken. They do not always increase, for some of the figures may be 0.

7. If the word "all" in the hypothesis of the axiom of continuity be assumed to refer only to rational numbers so that the statement becomes: If all rational numbers be divided into two classes $\cdots$, there shall be a number N (not necessarily rational) such that $\cdots$; then the conclusion may be taken as defining a number as the frontier of a sequence of rational numbers. Show that if two numbers X, Y be defined by two such sequences, and if the sum of the numbers be *defined* as the number defined by the sequence of the sums of corresponding terms as in Ex. 6, and if the product of the numbers be *defined* as the number defined by the sequence of the products as in Ex. 6, then the fundamental rules

$$X + Y = Y + X, \qquad XY = YX, \qquad (X + Y)Z = XZ + YZ$$

of arithmetic hold for the numbers X, Y, Z defined by sequences. In this way a complete theory of irrationals may be built up from the properties of rationals combined with the principle of continuity, namely, 1° by defining irrationals as frontiers of sequences of rationals, 2° by defining the operations of addition, multiplication, $\cdots$ as operations upon the rational numbers in the sequences, 3° by showing that the fundamental rules of arithmetic still hold for the irrationals.

8. Apply the principle of continuity to show that there is a positive number x such that $x^2 = 2$. To do this it should be shown that the rationals are divisible into two classes, those whose square is less than 2 and those whose square is not less than 2; and that these classes satisfy the requirements of the axiom of continuity. In like manner if a is any positive number and n is any positive integer, show that there is an x such that $x^n = a$.

21. Theorems on limits and on sets of points. The theorem on limits which is of fundamental algebraic importance is

THEOREM 2. If $R(x, y, z, \cdots)$ be any rational function of the variables $x, y, z, \cdots$, and if these variables are approaching limits $X, Y, Z, \cdots$, then the value of R approaches a limit and the limit is $R(X, Y, Z, \cdots)$, provided there is no division by zero.

As any rational expression is made up from its elements by combinations of addition, subtraction, multiplication, and division, it is sufficient to prove the theorem for these four operations. All except the last have been indicated in the above Ex. 3. As multiplication has been cared for, division need be considered only in the simple case of a reciprocal $1/x$. It must be proved that if $\lim x = X$, then $\lim (1/x) = 1/X$. Now

$$\left|\frac{1}{x} - \frac{1}{X}\right| = \frac{|x - X|}{|x||X|}, \qquad \text{by Ex. 3 } (\gamma) \text{ above.}$$

This quantity must be shown to be less than any assigned ϵ. As the quantity is complicated it will be replaced by a simpler one which is greater, owing to an increase in the denominator. Since $x \doteq X$, $x - X$ may be made numerically as small as desired, say less than ϵ', for all x's subsequent to some particular x. Hence if ϵ' be taken at least as small as $\frac{1}{2}|X|$, it appears that $|x|$ must be greater than $\frac{1}{2}|X|$. Then

$$\frac{|x - X|}{|x||X|} < \frac{|x - X|}{\frac{1}{2}|X|^2} = \frac{\epsilon'}{\frac{1}{2}|X|^2}, \qquad \text{by Ex. 3 } (\alpha) \text{ above,}$$

and if ϵ' be restricted to being less than $\frac{1}{2}|X|^2\epsilon$, the difference is less than ϵ and the theorem that $\lim (1/x) = 1/X$ is proved, and also Theorem 2. The necessity for the restriction $X \neq 0$ and the corresponding restriction in the statement of the theorem is obvious.

THEOREM 3. If when ϵ is given, no matter how small, it is possible to find a value of n so great that the difference $|x_{n+p} - x_n|$ between x_n and every subsequent term x_{n+p} in the suite $x_1, x_2, \cdots, x_n, \cdots$ is less than ϵ, the suite approaches a limit, and conversely.

The converse part has already been given as Ex. 1 above. The theorem itself is a consequence of the axiom of continuity. First note that as $|x_{n+p} - x_n| < \epsilon$ for all x's subsequent to x_n, the x's cannot become infinite. Suppose 1° that there is some number l such that no matter how remote x_n is in the suite, there are always subsequent values of x which are greater than l and others which are less than l. As all the x's after x_n lie in the interval 2ϵ and as l is less than some x's and greater than others, l must lie in that interval. Hence $|l - x_{n+p}| < 2\epsilon$ for all

x's subsequent to x_n. But now 2ϵ can be made as small as desired because ϵ can be taken as small as desired. Hence the definition of a limit applies and the x's approach l as a limit.

Suppose 2° that there is no such number l. Then every number k is such that either it is possible to go so far in the suite that all subsequent numbers x are as great as k or it is possible to go so far that all subsequent x's are less than k. Hence all numbers k are divided into two classes which satisfy the requirements of the axiom of continuity, and there must be a number N such that the x's ultimately come to lie between $N - \epsilon'$ and $N + \epsilon'$, no matter how small ϵ' is. Hence the x's approach N as a limit. Thus under either supposition the suite approaches a limit and the theorem is proved. It may be noted that under the second supposition the x's ultimately lie entirely upon one side of the point N and that the condition $|x_{n+p} - x_n| < \epsilon$ is not used except to show that the x's remain finite.

22. Consider next a set of points (or their correlative numbers) without any implication that they form a suite, that is, that one may be said to be subsequent to another. If there is only a finite number of points in the set, there is a point farthest to the right and one farthest to the left. If there is an infinity of points in the set, two possibilities arise. Either 1° it is not possible to assign a point K so far to the right that no point of the set is farther to the right — in which case the set is said to be *unlimited above* — or 2° there is a point K such that no point of the set is beyond K — and the set is said to be *limited above.* Similarly, a set may be *limited below* or *unlimited below.* If a set is limited above and below so that it is entirely contained in a finite interval, it is said merely to be *limited.* If there is a point C such that in any interval, no matter how small, surrounding C there are points of the set, then C is called a *point of condensation* of the set (C itself may or may not belong to the set).

THEOREM 4. Any infinite set of points which is limited has an upper frontier (maximum ?), a lower frontier (minimum ?), and at least one point of condensation.

Before proving this theorem, consider three infinite sets as illustrations:

(α) 1, 1.9, 1.99, 1.999, $\cdots$, (β) $-2, \cdots, -1.99, -1.9, -1$,

(γ) $-1, -\frac{1}{2}, -\frac{1}{4}, \cdots, \frac{1}{4}, \frac{1}{2}, 1$.

In (α) the element 1 is the minimum and serves also as the lower frontier; it is clearly not a point of condensation, but is isolated. There is no maximum; but 2 is the upper frontier and also a point of condensation. In (β) there is a maximum -1 and a minimum -2 (for -2 has been incorporated with the set). In (γ) there is a maximum and minimum; the point of condensation is 0. If one could be sure that an infinite set had a maximum and minimum, as is the case with finite sets, there would be no need of considering upper and lower frontiers. It is clear that if the upper or lower frontier belongs to the set, there is a maximum or minimum and the frontier is not necessarily a point of condensation; whereas *if the frontier does not belong to the set, it is necessarily a point of condensation and the corresponding extreme point is missing.*

To prove that there is an upper frontier, divide the points of the line into two classes, one consisting of points which are to the left of some point of the set, the other of points which are not to the left of any point of the set — then apply the axiom. Similarly for the lower frontier. To show the existence of a point of condensation, note that as there is an infinity of elements in the set, any point p is such that either there is an infinity of points of the set to the right of it or there is not. Hence the two classes into which all points are to be assorted are suggested, and the application of the axiom offers no difficulty.

EXERCISES

1. In a manner analogous to the proof of Theorem 2, show that

$$(\alpha)\ \lim_{x \doteq 0} \frac{x-1}{x-2} = \frac{1}{2}, \qquad (\beta)\ \lim_{x \doteq 2} \frac{3x-1}{x+5} = \frac{5}{7}, \qquad (\gamma)\ \lim_{x \doteq -1} \frac{x^2+1}{x^3-1} = -1.$$

2. Given an infinite series $S = u_1 + u_2 + u_3 + \cdots$. Construct the suite

$$S_1 = u_1,\ S_2 = u_1 + u_2,\ S_3 = u_1 + u_2 + u_3,\ \cdots,\ S_i = u_1 + u_2 + \cdots + u_i,\ \cdots,$$

where S_i is the sum of the first i terms. Show that Theorem 3 gives: The necessary and sufficient condition that the series S converge is that it is possible to find an n so large that $|S_{n+p} - S_n|$ shall be less than an assigned ϵ for all values of p. It is to be understood that a series *converges* when the suite of S's approaches a limit, and conversely.

3. If in a series $u_1 - u_2 + u_3 - u_4 + \cdots$ the terms approach the limit 0, are alternately positive and negative, and each term is less than the preceding, the series converges. Consider the suites $S_1, S_3, S_5, \cdots$ and $S_2, S_4, S_6, \cdots$.

4. Given three infinite suites of numbers

$$x_1, x_2, \cdots, x_n, \cdots; \qquad y_1, y_2, \cdots, y_n, \cdots; \qquad z_1, z_2, \cdots, z_n, \cdots$$

of which the first never decreases, the second never increases, and the terms of the third lie between corresponding terms of the first two, $x_n \leqq z_n \leqq y_n$. Show that the suite of z's has a point of condensation at or between the limits approached by the x's and by the y's; and that if $\lim x = \lim y = l$, then the z's approach l as a limit.

5. Restate the definitions and theorems on sets of points in arithmetic terms.

6. Give the details of the proof of Theorem 4. Show that the proof as outlined gives the least point of condensation. How would the proof be worded so as to give the greatest point of condensation? Show that if a set is limited above, it has an upper frontier but need not have a lower frontier.

7. If a set of points is such that between any two there is a third, the set is said to be *dense*. Show that the rationals form a dense set; also the irrationals. Show that any point of a dense set is a point of condensation for the set.

8. Show that the rationals p/q where $q < K$ do not form a dense set — in fact are a finite set in any limited interval. Hence in regarding any irrational as the limit of a set of rationals it is necessary that the denominators and also the numerators should become infinite.

9. Show that if an infinite set of points lies in a limited region of the plane, say in the rectangle $a \leqq x \leqq b$, $c \leqq y \leqq d$, there must be at least one point of condensation of the set. Give the necessary definitions and apply the axiom of continuity successively to the abscissas and ordinates.

23. Real functions of a real variable. *If x be a variable which takes on a certain set of values of which the totality may be denoted by $[x]$ and if y is a second variable the value of which is uniquely determined for each x of the set $[x]$, then y is said to be a function of x defined over the set $[x]$.* The terms "limited," "unlimited," "limited above," "unlimited below," $\cdots$ are applied to a function if they are applicable to the set $[y]$ of values of the function. Hence Theorem 4 has the corollary:

THEOREM 5. If a function is limited over the set $[x]$, it has an upper frontier M and a lower frontier m for that set.

If the function takes on its upper frontier M, that is, if there is a value x_0 in the set $[x]$ such that $f(x_0) = M$, the function has the absolute *maximum* M at x_0; and similarly with respect to the lower frontier. In any case, the difference $M - m$ between the upper and lower frontiers is called the *oscillation* of the function for the set $[x]$. The set $[x]$ is generally an interval.

Consider some illustrations of functions and sets over which they are defined. The reciprocal $1/x$ is defined for all values of x save 0. In the neighborhood of 0 the function is unlimited above for positive x's and unlimited below for negative x's. It should be noted that the function is not limited in the interval $0 < x \leqq a$ but is limited in the interval $\epsilon \leqq x \leqq a$ where ϵ is any assigned positive number. The function $+\sqrt{x}$ is defined for all positive x's including 0 and is limited below. It is not limited above for the totality of all positive numbers; but if K is assigned, the function is limited in the interval $0 \leqq x \leqq K$. The factorial function $x!$ is defined only for positive integers, is limited below by the value 1, but is not limited above unless the set $[x]$ is limited above. The function $E(x)$ denoting the integer not greater than x or "the integral part of x" is defined for all positive numbers — for instance $E(3) = E(\pi) = 3$. This function is not expressed, like the elementary functions of calculus, as a "formula"; it is defined by a definite law, however, and is just as much of a function as $x^2 + 3x + 2$ or $\frac{1}{2}\sin^2 2x + \log x$. Indeed it should be noted that the elementary functions themselves are in the first instance defined by definite laws and that it is not until after they have been made the subject of considerable study and have been largely developed along analytic lines that they appear as formulas. The ideas of function and formula are essentially distinct and the latter is essentially secondary to the former.

The definition of function as given above excludes the so-called *multiple-valued* functions such as $\sqrt{x}$ and $\sin^{-1} x$ where to a given value of x correspond more than one value of the function. It is usual, however, in treating multiple-valued functions to resolve the functions into different parts or *branches* so that each branch is a single-valued function. Thus $+\sqrt{x}$ is one branch and $-\sqrt{x}$ the other branch of $\sqrt{x}$; in fact when x is positive the symbol $\sqrt{x}$ is usually restricted to mean merely $+\sqrt{x}$ and thus becomes a single-valued symbol. One branch of $\sin^{-1}x$ consists of the values between $-\frac{1}{2}\pi$ and $+\frac{1}{2}\pi$, other branches give values between $\frac{1}{2}\pi$ and $\frac{3}{2}\pi$ or $-\frac{1}{2}\pi$ and $-\frac{3}{2}\pi$, and so on. Hence the term "function" will be restricted in this chapter to the single-valued functions allowed by the definition.

24. *If $x=a$ is any point of an interval over which $f(x)$ is defined, the function $f(x)$ is said to be continuous at the point $x=a$ if*

$$\lim_{x \doteq a} f(x) = f(a), \qquad \textit{no matter how } x \doteq a.$$

The function is said to be continuous in the interval if it is continuous at every point of the interval. If the function is not continuous at the point a, it is said to be *discontinuous* at a; and if it fails to be continuous at any one point of an interval, it is said to be discontinuous in the interval.

Theorem 6. If any finite number of functions are continuous (at a point or over an interval), any rational expression formed of those functions is continuous (at the point or over the interval) provided no division by zero is called for.

Theorem 7. If $y=f(x)$ is continuous at x_0 and takes the value $y_0=f(x_0)$ and if $z=\phi(y)$ is a continuous function of y at $y=y_0$, then $z=\phi[f(x)]$ will be a continuous function of x at x_0.

In regard to the definition of continuity note that a function cannot be continuous at a point unless it is defined at that point. Thus e^{-1/x^2} is not continuous at $x=0$ because division by 0 is impossible and the function is undefined. If, however, the function be defined at 0 as $f(0)=0$, the function becomes continuous at $x=0$. In like manner the function $1/x$ is not continuous at the origin, and in this case it is impossible to assign to $f(0)$ any value which will render the function continuous; the function becomes infinite at the origin and the very idea of becoming infinite precludes the possibility of approach to a definite limit. Again, the function $E(x)$ is in general continuous, but is discontinuous for integral values of x. When a function is discontinuous at $x=a$, the *amount of the discontinuity* is the limit of the oscillation $M-m$ of the function in the interval $a-\delta<x<a+\delta$ surrounding the point a when δ approaches zero as its limit. The discontinuity of $E(x)$ at each integral value of x is clearly 1; that of $1/x$ at the origin is infinite no matter what value is assigned to $f(0)$.

In case the interval over which $f(x)$ is defined has end points, say $a \leqq x \leqq b$, the question of continuity at $x=a$ must of course be decided by allowing x to approach a from the right-hand side only; and similarly it is a question of left-handed approach to b. In general, if for any reason it is desired to restrict the approach of a variable to its limit to being one-sided, the notations $x \doteq a^+$ and $x \doteq b^-$ respectively are used to denote approach through greater values (right-handed) and through lesser values (left-handed). It is not necessary to make this specification in the case of the ends of an interval; for it is understood that x shall take on only values in the interval in question. It should be noted that

$\lim f(x) = f(x_0)$ when $x \doteq x_0^+$ in no wise implies the continuity of $f(x)$ at x_0; a simple example is that of $E(x)$ at the positive integral points.

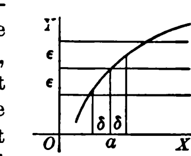
The proof of Theorem 6 is an immediate corollary application of Theorem 2. For

$$\lim R[f(x), \phi(x)\cdots] = R[\lim f(x), \lim \phi(x), \cdots] = R[f(\lim x), \phi(\lim x), \cdots],$$

and the proof of Theorem 7 is equally simple.

THEOREM 8. If $f(x)$ is continuous at $x = a$, then for any positive ϵ which has been assigned, no matter how small, there may be found a number δ such that $|f(x) - f(a)| < \epsilon$ in the interval $|x - a| < \delta$, and hence in this interval the oscillation of $f(x)$ is less than 2ϵ. And conversely, if these conditions hold, the function is continuous.

This theorem is in reality nothing but a restatement of the definition of continuity combined with the definition of a limit. For "$\lim f(x) = f(a)$ when $x \doteq a$, no matter how" means that the difference between $f(x)$ and $f(a)$ can be made as small as desired by taking x sufficiently near to a; and conversely. The reason for this restatement is that the present form is more amenable to analytic operations. It also suggests the geometric picture which corresponds to the usual idea of continuity in graphs. For the theorem states that if the two lines $y = f(a) \pm \epsilon$ be drawn, the graph of the function remains between them for at least the short distance δ on each side of $x = a$; and as ϵ may be assigned a value as small as desired, the graph cannot exhibit breaks. On the other hand it should be noted that the actual physical graph is not a curve but a band, a two-dimensional region of greater or less breadth, and that a function could be discontinuous at every point of an interval and yet lie entirely within the limits of any given physical graph.



It is clear that δ, which has to be determined *subsequently* to ϵ, is in general more and more restricted as ϵ is taken smaller and that for different points it is more restricted as the graph rises more rapidly. Thus if $f(x) = 1/x$ and $\epsilon = 1/1000$, δ can be nearly 1/10 if $x_0 = 100$, but must be slightly less than 1/1000 if $x_0 = 1$, and something less than 10^{-6} if x is 10^{-3}. Indeed, if x be allowed to approach zero, the value δ for any assigned ϵ also approaches zero; and although the function $f(x) = 1/x$ is continuous in the interval $0 < x \leqq 1$ and for any given x_0 and ϵ a number δ may be found such that $|f(x) - f(x_0)| < \epsilon$ when $|x - x_0| < \delta$, yet it is not possible to assign a number δ which shall serve *uniformly* for all values of x_0.

25. THEOREM 9. If a function $f(x)$ is continuous in an interval $a \leqq x \leqq b$ with end points, it is possible to find a δ such that $|f(x) - f(x_0)| < \epsilon$ when $|x - x_0| < \delta$ for all points x_0; and the function is said to be *uniformly continuous*.

The proof is conducted by the method of reductio ad absurdum. Suppose ϵ is assigned. Consider the suite of values $\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \cdots$, or any other suite which approaches zero as a limit. Suppose that no one of these values will serve as a δ for all points of the interval. Then there must be at least one point for which $\frac{1}{2}$ will not serve, at least one for which $\frac{1}{4}$ will not serve, at least one for which $\frac{1}{8}$ will not serve, and so on indefinitely. This infinite set of points must have at least one point of condensation C such that in any interval surrounding C there are points for which 2^{-k} will not serve as δ, no matter how large k. But now by hypothesis $f(x)$ is continuous at C and hence a number δ can be found such that $|f(x) - f(C)| < \frac{1}{2}\epsilon$ when $|x - x_0| < 2\delta$. The oscillation of $f(x)$ in the whole interval 4δ is less than ϵ. Now if x_0 be any point in the middle half of this interval, $|x_0 - C| < \delta$; and if x satisfies the relation $|x - x_0| < \delta$, it must still lie in the interval 4δ and the difference $|f(x) - f(x_0)| < \epsilon$, being surely not greater than the oscillation of f in the whole interval. Hence it is possible to surround C with an interval so small that the same δ will serve for any point of the interval. This contradicts the former conclusion, and hence the hypothesis upon which that conclusion was based must have been false and it must have been possible to find a δ which would serve for all points of the interval. The reason why the proof would not apply to a function like $1/x$ defined in the interval $0 < x \leqq 1$ lacking an end point is precisely that the point of condensation C would be 0, and at 0 the function is not continuous and $|f(x) - f(C)| < \frac{1}{2}\epsilon$, $|x - C| < 2\delta$ could not be satisfied.

THEOREM 10. If a function is continuous in a region which includes its end points, the function is limited.

THEOREM 11. If a function is continuous in an interval which includes its end points, the function takes on its upper frontier and has a maximum M; similarly it has a minimum m.

These are successive corollaries of Theorem 9. For let ϵ be assigned and let δ be determined so as to serve uniformly for all points of the interval. Divide the interval $b - a$ into n successive intervals of length δ or less. Then in each such interval f cannot increase by more than ϵ nor decrease by more than ϵ. Hence f will be contained between the values $f(a) + n\epsilon$ and $f(a) - n\epsilon$, and is limited. And $f(x)$ has an upper and a lower frontier in the interval. Next consider the rational function $1/(M - f)$ of f. By Theorem 6 this is continuous in the interval unless the denominator vanishes, and if continuous it is limited. This, however, is impossible for the reason that, as M is a frontier of values of f, the difference $M - f$ may be made as small as desired. Hence $1/(M - f)$ is not continuous and there must be some value of x for which $f = M$.

THEOREM 12. If $f(x)$ is continuous in the interval $a \leqq x \leqq b$ with end points and if $f(a)$ and $f(b)$ have opposite signs, there is at least one point ξ, $a < \xi < b$, in the interval for which the function vanishes. And whether $f(a)$ and $f(b)$ have opposite signs or not, there is a point ξ, $a < \xi < b$, such that $f(\xi) = \mu$, where μ is any value intermediate between the maximum and minimum of f in the interval.

For convenience suppose that $f(a) < 0$. Then in the neighborhood of $x = a$ the function will remain negative on account of its continuity; and in the neighborhood of b it will remain positive. Let ξ be the lower frontier of values of x which make $f(x)$ positive. Suppose that $f(\xi)$ were either positive or negative. Then as f is continuous, an interval could be chosen surrounding ξ and so small that f remained positive or negative in that interval. In neither case could ξ be the lower frontier of positive values. Hence the contradiction, and $f(\xi)$ must be zero. To prove the second part of the theorem, let c and d be the values of x which make f a minimum and maximum. Then the function $f-\mu$ has opposite signs at c and d, and must vanish at some point of the interval between c and d; and hence a fortiori at some point of the interval from a to b.

EXERCISES

1. Note that x is a continuous function of x, and that consequently it follows from Theorem 6 that any rational fraction $P(x)/Q(x)$, where P and Q are polynomials in x, must be continuous for all x's except roots of $Q(x)=0$.

2. Graph the function $x-E(x)$ for $x \geqq 0$ and show that it is continuous except for integral values of x. Show that it is limited, has a minimum 0, an upper frontier 1, but no maximum.

3. Suppose that $f(x)$ is defined for an infinite set $[x]$ of which $x=a$ is a point of condensation (not necessarily itself a point of the set). Suppose

$$\lim_{x', x'' \doteq a} [f(x')-f(x'')]=0 \quad \text{or} \quad |f(x')-f(x'')|<\epsilon,\ |x'-a|<\delta,\ |x''-a|<\delta,$$

when x' and x'' regarded as *independent* variables approach a as a limit (passing only over values of the set $[x]$, of course). Show that $f(x)$ approaches a limit as $x \doteq a$. By considering the set of values of $f(x)$, the method of Theorem 3 applies almost verbatim. Show that there is no essential change in the proof if it be assumed that x' and x'' become infinite, the set $[x]$ being unlimited instead of having a point of condensation a.

4. From the formula $\sin x < x$ and the formulas for $\sin u - \sin v$ and $\cos u - \cos v$ show that $\Delta \sin x$ and $\Delta \cos x$ are numerically less than $2|\Delta x|$; hence infer that $\sin x$ and $\cos x$ are continuous functions of x for all values of x.

5. What are the intervals of continuity for $\tan x$ and $\csc x$? If $\epsilon = 10^{-4}$, what are approximately the largest available values of δ that will make $|f(x)-f(x_0)|<\epsilon$ when $x_0 = 1°, 30°, 60°, 89°$ for each? Use a four-place table.

6. Let $f(x)$ be defined in the interval from 0 to 1 as equal to 0 when x is irrational and equal to $1/q$ when x is rational and expressed as a fraction p/q in lowest terms. Show that f is continuous for irrational values and discontinuous for rational values. Ex. 8, p. 39, will be of assistance in treating the irrational values.

7. Note that in the definition of continuity a generalization may be introduced by allowing the set $[x]$ over which f is defined to be any set each point of which is a point of condensation of the set, and that hence continuity over a dense set (Ex. 7 above), say the rationals or irrationals, may be defined. This is important because many functions are in the first instance defined only for rationals and are subsequently defined for irrationals by interpolation. Note that if a function is continuous over a dense set (say, the rationals), it does not follow that it is uniformly continuous over the set. For the point of condensation C which was used in the proof of Theorem 9 may not be a point of the set (may be irrational), and the proof would fall through for the same reason that it would in the case of $1/x$ in the interval $0<x \leqq 1$, namely, because it could not be affirmed that the function was continuous at C. Show that if a function is defined and is uniformly continuous over a dense set, the value $f(x)$ will approach a limit when x approaches any value a (not necessarily of the set, but situated between the upper and lower

frontiers of the set), and that if this limit be defined as the value of $f(a)$, the function will remain continuous. Ex. 3 may be used to advantage.

8. By factoring $(x + \Delta x)^n - x^n$, show for integral values of n that when $0 \leqq x \leqq K$, then $\Delta (x^n) < nK^{n-1} \Delta x$ for small Δx's and consequently x^n is uniformly continuous in the interval $0 \leqq x \leqq K$. If it be assumed that x^n has been defined only for rational x's, it follows from Ex. 7 that the definition may be extended to all x's and that the resulting x^n will be continuous.

9. Suppose (α) that $f(x) + f(y) = f(x + y)$ for any numbers x and y. Show that $f(n) = nf(1)$ and $nf(1/n) = f(1)$, and hence infer that $f(x) = xf(1) = Cx$, where $C = f(1)$, for all rational x's. From Ex. 7 it follows that if $f(x)$ is continuous, $f(x) = Cx$ for all x's. Consider (β) the function $f(x)$ such that $f(x)f(y) = f(x + y)$. Show that it is $Ce^x = a^x$.

10. Show by Theorem 12 that if $y = f(x)$ is a continuous constantly increasing function in the interval $a \leqq x \leqq b$, then to each value of y corresponds a single value of x so that the function $x = f^{-1}(y)$ exists and is single-valued; show also that it is continuous and constantly increasing. State the corresponding theorem if $f(x)$ is constantly decreasing. The function $f^{-1}(y)$ is called the *inverse* function to $f(x)$.

11. Apply Ex. 10 to discuss $y = \sqrt[n]{x}$, where n is integral, x is positive, and only positive roots are taken into consideration.

12. In arithmetic it may readily be shown that the equations

$$a^m a^n = a^{m+n}, \qquad (a^m)^n = a^{mn}, \qquad a^n b^n = (ab)^n,$$

are true when a and b are rational and positive and when m and n are any positive and negative integers or zero. (α) Can it be inferred that they hold when a and b are positive irrationals? (β) How about the extension of the fundamental inequalities

$$x^n > 1, \quad \text{when} \quad x > 1, \qquad x^n < 1, \quad \text{when} \quad 0 \leqq x < 1$$

to all rational values of n and the proof of the inequalities

$$x^m > x^n \quad \text{if} \quad m > n \quad \text{and} \quad x > 1, \qquad x^m < x^n \quad \text{if} \quad m > n \quad \text{and} \quad 0 < x < 1.$$

(γ) Next consider x as held constant and the exponent n as variable. Discuss the exponential function a^x from this relation, and Exs. 10, 11, and other theorems that may seem necessary. Treat the logarithm as the inverse of the exponential.

26. The derivative. *If $x = a$ is a point of an interval over which $f(x)$ is defined and if the quotient*

$$\frac{\Delta f}{\Delta x} = \frac{f(a + h) - f(h)}{h}, \qquad h = \Delta x,$$

approaches a limit when h approaches zero, no matter how, the function $f(x)$ is said to be differentiable at $x = a$ and the value of the limit of the quotient is the derivative $f'(a)$ of f at $x = a$. In the case of differentiability, the definition of a limit gives

$$\frac{f(a + h) - f(a)}{h} = f'(a) + \eta \quad \text{or} \quad f(a + h) - f(a) = hf'(a) + \eta h, \quad (1)$$

where $\lim \eta = 0$ when $\lim h = 0$, no matter how.

In other words if ϵ is given, a δ can be found so that $|\eta|<\epsilon$ when $|h|<\delta$. This shows that a function differentiable at a as in (1) is continuous at a. For

$$|f(a+h)-f(a)| \leqq |f'(a)|\delta + \epsilon\delta, \qquad |h|<\delta.$$

If the limit of the quotient exists when $h \doteq 0$ through positive values only, the function has a right-hand derivative which may be denoted by $f'(a^+)$ and similarly for the left-hand derivative $f'(a^-)$. At the end points of an interval the derivative is always considered as one-handed; but for interior points the right-hand and left-hand derivatives must be equal if the function is to have a derivative (unqualified). The function is said to have an *infinite derivative* at a if the quotient becomes infinite as $h \doteq 0$; but if a is an interior point, the quotient must become positively infinite or negatively infinite for all manners of approach and not positively infinite for some and negatively infinite for others. Geometrically this allows a vertical tangent with an inflection point, but not with a cusp as in Fig. 3, p. 8. If infinite derivatives are allowed, the function may have a derivative and yet be discontinuous, as is suggested by any figure where $f(a)$ is any value between $\lim f(x)$ when $x \doteq a^+$ and $\lim f(x)$ when $x \doteq a^-$.

Theorem 13. If a function takes on its maximum (or minimum) at an interior point of the interval of definition and if it is differentiable at that point, the derivative is zero.

Theorem 14. *Rolle's Theorem.* If a function $f(x)$ is continuous over an interval $a \leqq x \leqq b$ with end points and vanishes at the ends and has a derivative at each interior point $a<x<b$, there is some point ξ, $a<\xi<b$, such that $f'(\xi)=0$.

Theorem 15. *Theorem of the Mean.* If a function is continuous over an interval $a \leqq x \leqq b$ and has a derivative at each interior point, there is some point ξ such that

$$\frac{f(b)-f(a)}{b-a}=f'(\xi) \quad \text{or} \quad \frac{f(a+h)-f(a)}{h}=f'(a+\theta h),$$

where $h \leqq b-a$* and θ is a proper fraction, $0<\theta<1$.

To prove the first theorem, note that if $f(a)=M$, the difference $f(a+h)-f(a)$ cannot be negative for any value of h and the quotient $\Delta f/h$ cannot be positive when $h>0$ and cannot be negative when $h<0$. Hence the right-hand derivative cannot be positive and the left-hand derivative cannot be negative. As these two must be equal if the function has a derivative, it follows that they must be zero, and the derivative is zero. The second theorem is an immediate corollary. For as the function is continuous it must have a maximum and a minimum (Theorem 11) both of which cannot be zero unless the function is always zero in the interval. Now if the function is identically zero, the derivative is identically zero and the theorem is true; whereas if the function is not identically zero, either the maximum or minimum must be at an interior point, and at that point the derivative will vanish.

* That the theorem is true for any part of the interval from a to b if it is true for the whole interval follows from the fact that the conditions, namely, that f be continuous and that f' exist, hold for any part of the interval if they hold for the whole.

To prove the last theorem construct the auxiliary function

$$\psi(x) = f(x) - f(a) - (x-a)\frac{f(b)-f(a)}{b-a}, \qquad \psi'(x) = f'(x) - \frac{f(b)-f(a)}{b-a}.$$

As $\psi(a) = \psi(b) = 0$, Rolle's Theorem shows that there is some point for which $\psi'(\xi) = 0$, and if this value be substituted in the expression for $\psi'(x)$ the solution for $f'(\xi)$ gives the result demanded by the theorem. The proof, however, requires the use of the function $\psi(x)$ and its derivative and is not complete until it is shown that $\psi(x)$ really satisfies the conditions of Rolle's Theorem, namely, is continuous in the interval $a \leqq x \leqq b$ and has a derivative for every point $a < x < b$. The continuity is a consequence of Theorem 6; that the derivative exists follows from the direct application of the definition combined with the assumption that the derivative of f exists.

27. Theorem 16. If a function has a derivative which is identically zero in the interval $a \leqq x \leqq b$, the function is constant; and if two functions have derivatives equal throughout the interval, the functions differ by a constant.

Theorem 17. If $f(x)$ is differentiable and becomes infinite when $x \doteq a$, the derivative cannot remain finite as $x \doteq a$.

Theorem 18. If the derivative $f'(x)$ of a function exists and is a continuous function of x in the interval $a \leqq x \leqq b$, the quotient $\Delta f/h$ converges uniformly toward its limit $f'(x)$.

These theorems are consequences of the Theorem of the Mean. For the first,

$$f(a+h) - f(a) = hf'(a+\theta h) = 0, \quad \text{if} \quad h \leqq b - a, \quad \text{or} \quad f(a+h) = f(a).$$

Hence $f(x)$ is constant. And in case of two functions f and ϕ with equal derivatives, the difference $\psi(x) = f(x) - \phi(x)$ will have a derivative that is zero and the difference will be constant. For the second, let x_0 be a fixed value near a and suppose that in the interval from x_0 to a the derivative remained finite, say less than K. Then

$$|f(x_0+h) - f(x_0)| = |hf'(x_0+\theta h)| \leqq |h| K.$$

Now let $x_0 + h$ approach a and note that the left-hand term becomes infinite and the supposition that f' remained finite is contradicted. For the third, note that f', being continuous, must be uniformly continuous (Theorem 9), and hence that if ϵ is given, a δ may be found such that

$$\left|\frac{f(x+h)-f(x)}{h} - f'(x)\right| \leqq |f'(x+\theta h) - f'(x)| < \epsilon$$

when $|h| < \delta$ and for all x's in the interval; and the theorem is proved.

Concerning derivatives of higher order no special remarks are necessary. Each is the derivative of a definite function — the previous derivative. If the derivatives of the first n orders exist and are continuous, the derivative of order $n+1$ may or may not exist. In practical applications, however, the functions are generally indefinitely differentiable except at certain isolated points. The proof of Leibniz's Theorem (§ 8) may be revised so as to depend on elementary processes. Let the formula be assumed for a given value of n. The only terms which can contribute to the term $D^iuD^{n+1-i}v$ in the formula for the $(n+1)$st derivative of uv are the terms

$$\frac{n(n-1)\cdots(n-i+2)}{1\cdot 2\cdots(i-1)}D^{i-1}u\,D^{n+1-i}v, \qquad \frac{n(n-1)\cdots(n-i+1)}{1\cdot 2\cdots i}D^iu\,D^{n-i}v,$$

in which the first factor is to be differentiated in the first and the second in the second. The sum of the coefficients obtained by differentiating is

$$\frac{n(n-1)\cdots(n-i+2)}{1\cdot 2\cdots(i-1)}+\frac{n(n-1)\cdots(n-i+1)}{1\cdot 2\cdots i}=\frac{(n+1)n\cdots(n-i+2)}{1\cdot 2\cdots i},$$

which is precisely the proper coefficient for the term $D^iuD^{n+1-i}v$ in the expansion of the $(n+1)$st derivative of uv by Leibniz's Theorem.

With regard to this rule and the other elementary rules of operation (4)–(7) of the previous chapter it should be remarked that *a theorem* as well as a rule is involved — thus: If two functions u and v are differentiable at x_0, then the product uv is differentiable at x_0, and the value of the derivative is $u(x_0)v'(x_0)+u'(x_0)v(x_0)$. And similar theorems arise in connection with the other rules. As a matter of fact the ordinary proof needs only to be gone over with care in order to convert it into a rigorous demonstration. But care does need to be exercised both in stating the theorem and in looking to the proof. For instance, the above theorem concerning a product is not true if infinite derivatives are allowed. For let u be -1, 0, or $+1$ according as x is negative, 0, or positive, and let $v=x$. Now v has always a derivative which is 1 and u has always a derivative which is 0, $+\infty$, or 0 according as x is negative, 0, or positive. The product uv is $|x|$, of which the derivative is -1 for negative x's, $+1$ for positive x's, and *nonexistent* for 0. Here the product has no derivative at 0, although each factor has a derivative, and it would be useless to have a formula for attempting to evaluate something that did not exist.

EXERCISES

1. Show that if at a point the derivative of a function exists and is positive, the function must be increasing at that point.

2. Suppose that the derivatives $f'(a)$ and $f'(b)$ exist and are not zero. Show that $f(a)$ and $f(b)$ are relative maxima or minima of f in the interval $a\leqq x\leqq b$, and determine the precise criteria in terms of the signs of the derivatives $f'(a)$ and $f'(b)$.

3. Show that if a continuous function has a positive right-hand derivative at every point of the interval $a\leqq x\leqq b$, then $f(b)$ is the maximum value of f. Similarly, if the right-hand derivative is negative, show that $f(b)$ is the minimum of f.

4. Apply the Theorem of the Mean to show that if $f'(x)$ is continuous at a, then

$$\lim_{x',\,x''\doteq a}\frac{f(x')-f(x'')}{x'-x''}=f'(a),$$

x' and x'' being regarded as independent.

5. Form the increments of a function f for *equicrescent* values of the variable:

$$\Delta_1 f=f(a+h)-f(a), \qquad \Delta_2 f=f(a+2h)-f(a+h),$$
$$\Delta_3 f=f(a+3h)-f(a+2h),\cdots.$$

These are called first differences; the differences of these differences are

$$\Delta_1^2 f = f(a+2h) - 2f(a+h) + f(a),$$
$$\Delta_2^2 f = f(a+3h) - 2f(a+2h) + f(a+h), \cdots$$

which are called the second differences; in like manner there are third differences

$$\Delta_1^3 f = f(a+3h) - 3f(a+2h) + 3f(a+h) - f(a), \cdots$$

and so on. Apply the Law of the Mean to all the differences and show that

$$\Delta_1^2 f = h^2 f''(a+\theta_1 h+\theta_2 h), \qquad \Delta_1^3 f = h^3 f'''(a+\theta_1 h+\theta_2 h+\theta_3 h), \cdots.$$

Hence show that if the first n derivatives of f are continuous at a, then

$$f''(a) = \lim_{h \doteq 0} \frac{\Delta^2 f}{h^2}, \qquad f'''(a) = \lim_{h \doteq 0} \frac{\Delta^3 f}{h^3}, \qquad \cdots, \qquad f^{(n)}(a) = \lim_{h \doteq 0} \frac{\Delta^n f}{h^n}.$$

6. *Cauchy's Theorem.* If $f(x)$ and $\phi(x)$ are continuous over $a \leqq x \leqq b$, have derivatives at each interior point, and if $\phi'(x)$ does not vanish in the interval,

$$\frac{f(b)-f(a)}{\phi(b)-\phi(a)} = \frac{f'(\xi)}{\phi'(\xi)} \quad \text{or} \quad \frac{f(a+h)-f(a)}{\phi(a+h)-\phi(a)} = \frac{f'(a+\theta h)}{\phi'(a+\theta h)}.$$

Prove that this follows from the application of Rolle's Theorem to the function

$$\psi(x) = f(x) - f(a) - [\phi(x) - \phi(a)] \frac{f(b)-f(a)}{\phi(b)-\phi(a)}.$$

7. One application of Ex. 6 is to the theory of indeterminate forms. Show that if $f(a) = \phi(a) = 0$ and if $f'(x)/\phi'(x)$ approaches a limit when $x \doteq a$, then $f(x)/\phi(x)$ will approach the same limit.

8. *Taylor's Theorem.* Note that the form $f(b) = f(a) + (b-a)f'(\xi)$ is one way of writing the Theorem of the Mean. By the application of Rolle's Theorem to

$$\psi(x) = f(b) - f(x) - (b-x)f'(x) - (b-x)^2 \frac{f(b) - f(a) - (b-a)f'(a)}{(b-a)^2},$$

show
$$f(b) = f(a) + (b-a)f'(a) + \frac{(b-a)^2}{2} f''(\xi),$$

and to
$$\psi(x) = f(b) - f(x) - (b-x)f'(x) - \frac{(b-x)^2}{2} f''(x) - \cdots - \frac{(b-x)^{n-1}}{(n-1)!} f^{(n-1)}(x)$$
$$- \frac{(b-x)^n}{(b-a)^n}\Big[f(b) - f(a) - (b-a)f'(a) - \frac{(b-a)^2}{2} f''(a) - \cdots - \frac{(b-a)^{n-1}}{(n-1)!} f^{(n-1)}(a) \Big],$$

show
$$f(b) = f(a) + (b-a)f'(a) + \frac{(b-a)^2}{2} f''(a) + \cdots + \frac{(b-a)^{n-1}}{(n-1)!} f^{(n-1)}(a) + \frac{(b-a)^n}{n!} f^{(n)}(\xi).$$

What are the restrictions that must be imposed on the function and its derivatives?

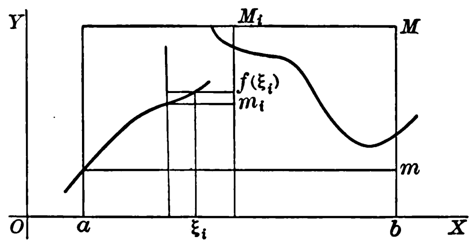
9. If a continuous function over $a \leqq x \leqq b$ has a right-hand derivative at each point of the interval which is zero, show that the function is constant. Apply Ex. 2 to the functions $f(x) + \epsilon(x-a)$ and $f(x) - \epsilon(x-a)$ to show that the maximum difference between the functions is $2\epsilon(b-a)$ and that f must therefore be constant.

10. State and prove the theorems implied in the formulas (4)–(6), p. 2.

11. Consider the extension of Ex. 7, p. 44, to derivatives of functions defined over a dense set. If the derivative exists and is uniformly continuous over the dense set, what of the existence and continuity of the derivative of the function when its definition is extended as there indicated?

12. If $f(x)$ has a finite derivative at each point of the interval $a \leqq x \leqq b$, the derivative $f'(x)$ must take on every value intermediate between any two of its values. To show this, take first the case where $f'(a)$ and $f'(b)$ have opposite signs and show, by the continuity of f and by Theorem 13 and Ex. 2, that $f'(\xi) = 0$. Next if $f'(a) < \mu < f'(b)$ without any restrictions on $f'(a)$ and $f'(b)$, consider the function $f(x) - \mu x$ and its derivative $f'(x) - \mu$. Finally, prove the complete theorem. It should be noted that the continuity of $f'(x)$ is not assumed, nor is it proved; for there are functions which take every value intermediate between two given values and yet are not continuous.

28. Summation and integration. Let $f(x)$ be defined and limited over the interval $a \leqq x \leqq b$ and let M, m, and $O = M - m$ be the upper frontier, lower frontier, and oscillation of $f(x)$ in the interval. Let $n - 1$ points of division be introduced in the interval dividing it into n consecutive intervals $\delta_1, \delta_2, \cdots, \delta_n$ of which the largest has the length Δ and let M_i, m_i, O_i, and $f(\xi_i)$ be the upper and lower frontiers, the oscillation, and any value of the function in the interval δ_i. Then the inequalities



$$m\delta_i \leqq m_i\delta_i \leqq f(\xi_i)\,\delta_i \leqq M_i\delta_i \leqq M\delta_i$$

will hold, and if these terms be summed up for all n intervals,

$$m\,(b-a) \leqq \sum m_i\delta_i \leqq \sum f(\xi_i)\,\delta_i \leqq \sum M_i\delta_i \leqq M\,(b-a) \qquad (A)$$

will also hold. Let $s = \Sigma m_i\delta_i$, $\sigma = \Sigma f(\xi_i)\,\delta_i$, and $S = \Sigma M_i\delta_i$. From (A) it is clear that the difference $S - s$ does not exceed

$$(M - m)\,(b - a) = O\,(b - a),$$

the product of the length of the interval by the oscillation in it. The values of the sums S, s, σ will evidently depend on the number of parts into which the interval is divided and on the way in which it is divided into that number of parts.

THEOREM 19. If n' additional points of division be introduced into the interval, the sum S' constructed for the $n + n' - 1$ points of division cannot be greater than S and cannot be less than S by more than $n'O\Delta$. Similarly, s' cannot be less than s and cannot exceed s by more than $n'O\Delta$.

.THEOREM 20. There exists a lower frontier L for all possible methods of constructing the sum S and an upper frontier l for s.

THEOREM 21. *Darboux's Theorem.* When ϵ is assigned it is possible to find a Δ so small that for all methods of division for which $\delta_i \leqq \Delta$, the sums S and s shall differ from their frontier values L and l by less than any preassigned ϵ.

To prove the first theorem note that although (A) is written for the whole interval from a to b and for the sums constructed on it, yet it applies equally to any part of the interval and to the sums constructed on that part. Hence if $S_i = M_i\delta_i$ be the part of S due to the interval δ_i and if S_i' be the part of S' due to this interval after the introduction of some of the additional points into it, $m_i\delta_i \leqq S_i' \leqq S_i = M_i\delta_i$. Hence S_i' is not greater than S_i (and as this is true for each interval δ_i, S' is not greater than S) and, moreover, $S_i - S_i'$ is not greater than $O_i\delta_i$ and a fortiori not greater than $O\Delta$. As there are only n' new points, not more than n' of the intervals δ_i can be affected, and hence the total decrease $S - S'$ in S cannot be more than $n'O\Delta$. The treatment of s is analogous.

Inasmuch as (A) shows that the sums S and s are limited, it follows from Theorem 4 that they possess the frontiers required in Theorem 20. To prove Theorem 21 note first that as L is a frontier for all the sums S, there is some particular sum S which differs from L by as little as desired, say $\frac{1}{2}\epsilon$. For this S let n be the number of divisions. Now consider S' as any sum for which each δ_i is less than $\Delta = \frac{1}{2}\epsilon/nO$. If the sum S'' be constructed by adding the n points of division for S to the points of division for S', S'' cannot be greater than S and hence cannot differ from L by so much as $\frac{1}{2}\epsilon$. Also S'' cannot be greater than S' and cannot be less than S' by more than $nO\Delta$, which is $\frac{1}{2}\epsilon$. As S'' differs from L by less than $\frac{1}{2}\epsilon$ and S' differs from S'' by less than $\frac{1}{2}\epsilon$, S' cannot differ from L by more than ϵ, which was to be proved. The treatment of s and l is analogous.

29. If indices are introduced to indicate the interval for which the frontiers L and l are calculated and if β lies in the interval from a to b, then L_a^β and l_a^β will be functions of β.

THEOREM 22. The equations $L_a^b = L_a^c + L_c^b$, $a < c < b$; $L_a^b = -L_b^a$; $L_a^b = \mu(b-a)$, $m \leqq \mu \leqq M$, hold for L, and similar equations for l. As functions of β, L_a^β and l_a^β are continuous, and if $f(x)$ is continuous, they are differentiable and have the common derivative $f(\beta)$.

To prove that $L_a^b = L_a^c + L_c^b$, consider c as one of the points of division of the interval from a to b. Then the sums S will satisfy $S_a^b = S_a^c + S_c^b$, and as the limit of a sum is the sum of the limits, the corresponding relation must hold for the frontier L. To show that $L_a^b = -L_b^a$ it is merely necessary to note that $S_a^b = -S_b^a$ because in passing from b to a the intervals δ_i must be taken with the sign opposite to that which they have when the direction is from a to b. From (A) it appears that $m(b-a) \leqq S_a^b \leqq M(b-a)$ and hence in the limit $m(b-a) \leqq L_a^b \leqq M(b-a)$.

Hence there is a value μ, $m \leqq \mu \leqq M$, such that $L_a^b = \mu(b-a)$. To show that L_a^β is a continuous function of β, take $K > |M|$ and $|m|$, and consider the relations

$$L_a^{\beta+h} - L_a^\beta = L_a^\beta + L_\beta^{\beta+h} - L_a^\beta = L_\beta^{\beta+h} = \mu h, \qquad |\mu| < K,$$
$$L_a^{\beta-h} - L_a^\beta = L_a^{\beta-h} - L_a^{\beta-h} - L_{\beta-h}^\beta = - L_{\beta-h}^\beta = - \mu' h, \qquad |\mu'| < K.$$

Hence if ϵ is assigned, a δ may be found, namely $\delta < \epsilon/K$, so that $|L_a^{\beta\pm h} - L_a^\beta| < \epsilon$ when $h < \delta$ and L_a^β is therefore continuous. Finally consider the quotients

$$\frac{L_a^{\beta+h} - L_a^\beta}{h} = \mu \quad \text{and} \quad \frac{L_a^{\beta-h} - L_a^\beta}{-h} = \mu',$$

where μ is some number between the maximum and minimum of $f(x)$ in the interval $\beta \leqq x \leqq \beta + h$ and, if f is continuous, is some value $f(\xi)$ of f in that interval and where $\mu' = f(\xi')$ is some value of f in the interval $\beta - h \leqq x \leqq \beta$. Now let $h \doteq 0$. As the function f is continuous, $\lim f(\xi) = f(\beta)$ and $\lim f(\xi') = f(\beta)$. Hence the right-hand and left-hand derivatives exist and are equal and the function L_a^β has the derivative $f(\beta)$. The treatment of l is analogous.

THEOREM 23. For a given interval and function f, the quantities l and L satisfy the relation $l \leqq L$; and the necessary and sufficient condition that $L = l$ is that there shall be some division of the interval which shall make $\Sigma(M_i - m_i)\delta_i = \Sigma O_i\delta_i < \epsilon$.

If $L_a^b = l_a^b$, the function f is said to be integrable over the interval from a to b and the integral $\int_a^b f(x)\,dx$ is defined as the common value $L_a^b = l_a^b$. Thus the definite integral is defined.

THEOREM 24. If a function is integrable over an interval, it is integrable over any part of the interval and the equations

$$\int_a^c f(x)\,dx + \int_c^b f(x)\,dx = \int_a^b f(x)\,dx,$$

$$\int_a^b f(x)\,dx = -\int_b^a f(x)\,dx, \qquad \int_a^b f(x)\,dx = \mu(b-a)$$

hold; moreover, $\int_a^\beta f(x)\,dx = F(\beta)$ is a continuous function of β; and if $f(x)$ is continuous, the derivative $F'(\beta)$ will exist and be $f(\beta)$.

By (A) the sums S and s constructed for the same division of the interval satisfy the relation $S - s \geqq 0$. By Darboux's Theorem the sums S and s will approach the values L and l when the divisions are indefinitely decreased. Hence $L - l \geqq 0$. Now if $L = l$ and a Δ be found so that when $\delta_i < \Delta$ the inequalities $S - L < \frac{1}{2}\epsilon$ and $l - s < \frac{1}{2}\epsilon$ hold, then $S - s = \Sigma(M_i - m_i)\delta_i = \Sigma O_i\delta_i < \epsilon$; and hence the condition $\Sigma O_i\delta_i < \epsilon$ is seen to be necessary. Conversely if there is any method of division such that $\Sigma O_i\delta_i < \epsilon$, then $S - s < \epsilon$ and the lesser quantity $L - l$ must also be less than ϵ. But if the difference between two constant quantities can be made less than ϵ, where ϵ is arbitrarily assigned, the constant quantities are equal; and hence the condition is seen to be also sufficient. To show that if a function is integrable over an interval, it is integrable over any part of the interval, it is merely necessary to show that if $L_a^b = l_a^b$, then $L_\alpha^\beta = l_\alpha^\beta$ where α and β are two points of the interval. Here the condition $\Sigma O_i\delta_i < \epsilon$ applies; for if $\Sigma O_i\delta_i$ can be made less than ϵ for the whole interval, its value for any part of the interval, being less than for the whole, must be less than ϵ. The rest of Theorem 24 is a corollary of Theorem 22.

30.* THEOREM 25. A function is integrable over the interval $a \leqq x \leqq b$ if it is continuous in that interval.

THEOREM 26. If the interval $a \leqq x \leqq b$ over which $f(x)$ is defined and limited contains only a finite number of points at which f is discontinuous or if it contains an infinite number of points at which f is discontinuous but these points have only a finite number of points of condensation, the function is integrable.

THEOREM 27. If $f(x)$ is integrable over the interval $a \leqq x \leqq b$, the sum $\sigma = \Sigma f(\xi_i)\delta_i$ will approach the limit $\int_a^b f(x)\,dx$ when the individual intervals δ_i approach the limit zero, it being immaterial how they approach that limit or how the points ξ_i are selected in their respective intervals δ_i.

THEOREM 28. If $f(x)$ is continuous in an interval $a \leqq x \leqq b$, then $f(x)$ has an indefinite integral, namely $\int_a^x f(x)\,dx$, in the interval.

Theorem 25 may be reduced to Theorem 23. For as the function is continuous, it is possible to find a Δ so small that the oscillation of the function in any interval of length Δ shall be as small as desired (Theorem 9). Suppose Δ be chosen so that the oscillation is less than $\epsilon/(b-a)$. Then $\Sigma O_i\delta_i < \epsilon$ when $\delta_i < \Delta$; and the function is integrable. To prove Theorem 26, take first the case of a finite number of discontinuities. Cut out the discontinuities surrounding each value of x at which f is discontinuous by an interval of length δ. As the oscillation in each of these intervals is not greater than O, the contribution of these intervals to the sum $\Sigma O_i\delta_i$ is not greater than $On\delta$, where n is the number of the discontinuities. By taking δ small enough this may be made as small as desired, say less than $\frac{1}{2}\epsilon$. Now in each of the remaining parts of the interval $a \leqq x \leqq b$, the function f is continuous and hence integrable, and consequently the value of $\Sigma O_i\delta_i$ for these portions may be made as small as desired, say $\frac{1}{2}\epsilon$. Thus the sum $\Sigma O_i\delta_i$ for the whole interval can be made as small as desired and $f(x)$ is integrable. When there are points of condensation they may be treated just as the isolated points of discontinuity were treated. After they have been surrounded by intervals, there will remain over only a finite number of discontinuities. Further details will be left to the reader.

For the proof of Theorem 27, appeal may be taken to the fundamental relation (A) which shows that $s \leqq \sigma \leqq S$. Now let the number of divisions increase indefinitely and each division become indefinitely small. As the function is integrable, S and s approach the same limit $\int_a^b f(x)\,dx$, and consequently σ which is included between them must approach that limit. Theorem 28 is a corollary of Theorem 24 which states that as $f(x)$ is continuous, the derivative of $\int_a^x f(x)\,dx$ is $f(x)$. By definition, the indefinite integral is any function whose derivative is the integrand. Hence $\int_a^x f(x)\,dx$ is an indefinite integral of $f(x)$, and any other may be obtained by adding to this an arbitrary constant (Theorem 16). Thus it is seen that the proof of the existence of the indefinite integral for any given continuous function is made to depend on the theory of definite integrals.

EXERCISES

1. Rework some of the proofs in the text with l replacing L.

2. Show that the L obtained from $Cf(x)$, where C is a constant, is C times the L obtained from f. Also if u, v, w are all limited in the interval $a \leqq x \leqq b$, the L for the combination $u + v - w$ will be $L(u) + L(v) - L(w)$, where $L(u)$ denotes the L for u, etc. State and prove the corresponding theorems for definite integrals and hence the corresponding theorems for indefinite integrals.

3. Show that $\Sigma O_i \delta_i$ can be made less than an assigned ϵ in the case of the function of Ex. 6, p. 44. Note that $l = 0$, and hence infer that the function is integrable and the integral is zero. The proof may be made to depend on the fact that there are only a finite number of values of the function greater than any assigned value.

4. State with care and prove the results of Exs. 3 and 5, p. 29. What restriction is to be placed on $f(x)$ if $f(\xi)$ may replace μ?

5. State with care and prove the results of Ex. 4, p. 29, and Ex. 13, p. 30.

6. If a function is limited in the interval $a \leqq x \leqq b$ and never decreases, show that the function is integrable. This follows from the fact that $\Sigma O_i \leqq O$ is finite.

7. More generally, let $f(x)$ be such a function that ΣO_i remains less than some number K, no matter how the interval be divided. Show that f is integrable. Such a function is called a *function of limited variation* (§ 127).

8. *Change of variable.* Let $f(x)$ be continuous over $a \leqq x \leqq b$. Change the variable to $x = \phi(t)$, where it is supposed that $a = \phi(t_1)$ and $b = \phi(t_2)$, and that $\phi(t)$, $\phi'(t)$, and $f[\phi(t)]$ are continuous in t over $t_1 \leqq t \leqq t_2$. Show that

$$\int_a^b f(x)\,dx = \int_{t_1}^{t_2} f[\phi(t)]\,\phi'(t)\,dt \quad \text{or} \quad \int_{\phi(t_1)}^{\phi(t)} f(x)\,dx = \int_{t_1}^{t} f[\phi(t)]\,\phi'(t)\,dt.$$

Do this by showing that the derivatives of the two sides of the last equation with respect to t exist and are equal over $t_1 \leqq t \leqq t_2$, that the two sides vanish when $t = t_1$ and are equal, and hence that they must be equal throughout the interval.

9. *Osgood's Theorem.* Let α_i be a set of quantities which differ uniformly from $f(\xi_i)\,\delta_i$ by an amount $\zeta_i \delta_i$, that is, suppose

$$\alpha_i = f(\xi_i)\,\delta_i + \zeta_i \delta_i, \quad \text{where} \quad |\zeta_i| < \epsilon \quad \text{and} \quad a \leqq \xi \leqq b.$$

Prove that if f is integrable, the sum $\Sigma\alpha_i$ approaches a limit when $\delta_i \doteq 0$ and that the limit of the sum is $\int_a^b f(x)\,dx$.

10. Apply Ex. 9 to the case $\Delta f = f'\Delta x + \zeta\Delta x$ where f' is continuous to show directly that $f(b) - f(a) = \int_a^b f'(x)\,dx$. Also by regarding $\Delta x = \phi'(t)\,\Delta t + \zeta\Delta t$, apply to Ex. 8 to prove the rule for change of variable.

PART I. DIFFERENTIAL CALCULUS

CHAPTER III

TAYLOR'S FORMULA AND ALLIED TOPICS

31. Taylor's Formula. The object of Taylor's Formula is to express the value of a function $f(x)$ in terms of the values of the function and its derivatives at some one point $x = a$. Thus

$$f(x) = f(a) + (x-a)f'(a) + \frac{(x-a)^2}{2!}f''(a) + \cdots + \frac{(x-a)^{n-1}}{(n-1)!}f^{(n-1)}(a) + R. \qquad (1)$$

Such an expansion is necessarily true because the remainder R may be considered as defined by the equation; the real significance of the formula must therefore lie in the possibility of finding a simple expression for R, and there are several.

THEOREM. On the hypothesis that $f(x)$ and its first n derivatives exist and are continuous over the interval $a \leqq x \leqq b$, the function may be expanded in that interval into a polynomial in $x - a$,

$$f(x) = f(a) + (x-a)f'(a) + \frac{(x-a)^2}{2!}f''(a) + \cdots + \frac{(x-a)^{n-1}}{(n-1)!}f^{(n-1)}(a) + R, \qquad (1)$$

with the remainder R expressible in any one of the forms

$$R = \frac{(x-a)^n}{n!}f^{(n)}(\xi) = \frac{h^n(1-\theta)^{n-1}}{(n-1)!}f^{(n)}(\xi) = \frac{1}{(n-1)!}\int_0^h t^{n-1}f^{(n)}(a+h-t)\,dt, \qquad (2)$$

where $h = x - a$ and $a < \xi < x$ or $\xi = a + \theta h$ where $0 < \theta < 1$.

A first proof may be made to depend on Rolle's Theorem as indicated in Ex. 8, p. 49. Let x be regarded for the moment as constant, say equal to b. Construct the function $\psi(x)$ there indicated. Note that $\psi(a) = \psi(b) = 0$ and that the derivative $\psi'(x)$ is merely

$$\psi'(x) = -\frac{(b-x)^{n-1}}{(n-1)!}f^{(n)}(x) + n\frac{(b-x)^{n-1}}{(b-a)^n}\Big[f(b) - f(a) - (b-a)f'(a) - \cdots - \frac{(b-a)^{n-1}}{(n-1)!}f^{(n-1)}(a)\Big].$$

By Rolle's Theorem $\psi'(\xi) = 0$. Hence if ξ be substituted above, the result is

$$f(b) = f(a) + (b-a)f'(a) + \cdots + \frac{(b-a)^{n-1}}{(n-1)!}f^{(n-1)}(a) + \frac{(b-a)^n}{n!}f^{(n)}(\xi),$$

after striking out the factor $-(b-\xi)^{n-1}$, multiplying by $(b-a)^n/n$, and transposing $f(b)$. The theorem is therefore proved with the first form of the remainder. *This proof does not require the continuity of the nth derivative nor its existence at a and at b.*

The second form of the remainder may be found by applying Rolle's Theorem to

$$\psi(x) = f(b) - f(x) - (b-x)f'(x) - \cdots - \frac{(b-x)^{n-1}}{(n-1)!}f^{(n-1)}(x) - (b-x)P,$$

where P is determined so that $R = (b-a)P$. Note that $\psi(b) = 0$ and that by Taylor's Formula $\psi(a) = 0$. Now

$$\psi'(x) = -\frac{(b-x)^{n-1}}{(n-1)!}f^{(n)}(x) + P \quad \text{or} \quad P = f^{(n)}(\xi)\frac{(b-\xi)^{n-1}}{(n-1)!} \quad \text{since} \quad \psi'(\xi) = 0.$$

Hence if ξ be written $\xi = a + \theta h$ where $h = b - a$, then $b - \xi = b - a - \theta h = (b-a)(1-\theta)$.

And $$R = (b-a)P = (b-a)\frac{(b-a)^{n-1}(1-\theta)^{n-1}}{(n-1)!}f^{(n)}(\xi) = \frac{(b-a)^n(1-\theta)^{n-1}}{(n-1)!}f^{(n)}(\xi).$$

The second form of R is thus found. In this work as before, the result is proved for $x = b$, the end point of the interval $a \leqq x \leqq b$. But as the interval could be considered as terminating at any of its points, the proof clearly applies to any x in the interval.

A second proof of Taylor's Formula, and the easiest to remember, consists in integrating the nth derivative n times from a to x. The successive results are

$$\int_a^x f^{(n)}(x)\,dx = f^{n-1}(x)\Big]_a^x = f^{(n-1)}(x) - f^{(n-1)}(a).$$

$$\int_a^x\int_a^x f^{(n)}(x)\,dx^2 = \int_a^x f^{(n-1)}(x)\,dx - \int_a^x f^{(n-1)}(a)\,dx$$
$$= f^{(n-2)}(x) - f^{(n-2)}(a) - (x-a)f^{(n-1)}(a).$$

$$\int_a^x\int_a^x\int_a^x f^{(n)}(x)\,dx^3 = f^{(n-3)}(x) - f^{(n-3)}(a) - (x-a)f^{(n-2)}(a) - \frac{(x-a)^2}{2!}f^{(n-1)}(a).$$

$$\int_a^x\cdots\int_a^x f^{(n)}(x)\,dx^n = f(x) - f(a) - (x-a)f'(a) - \frac{(x-a)^2}{2!}f''(a) - \cdots - \frac{(x-a)^{n-1}}{(n-1)!}f^{(n-1)}(a).$$

The formula is therefore proved with R in the form $\int_a^x\cdots\int_a^x f^{(n)}(x)\,dx^n$. To transform this to the ordinary form, the Law of the Mean may be applied ((65), § 16). For

$$m(x-a) < \int_a^x f^{(n)}(x)\,dx < M(x-a), \qquad m\frac{(x-a)^n}{n!} < \int_a^x\cdots\int_a^x f^{(n)}(x)\,dx^n < M\frac{(x-a)^n}{n!},$$

where m is the least and M the greatest value of $f^{(n)}(x)$ from a to x. There is then some intermediate value $f^{(n)}(\xi)=\mu$ such that

$$\int_a^x\cdots\int_a^x f^{(n)}(x)\,dx^n=\frac{(x-a)^n}{n!}f^{(n)}(\xi).$$

This proof requires that the nth derivative be continuous and is less general.

The third proof is obtained by applying successive integrations by parts to the obvious identity $f(a+h)-f(a)=\int_0^h f'(a+h-t)\,dt$ to make the integrand contain higher derivatives.

$$f(a+h)-f(a)=\int_0^h f'(a+h-t)\,dt=tf'(a+h-t)\Big]_0^h+\int_0^h tf''(a+h-t)\,dt$$

$$=hf'(a)+\tfrac{1}{2}t^2f''(a+h-t)\Big]_0^h+\int_0^h\tfrac{1}{2}t^2f'''(a+h-t)\,dt$$

$$=hf'(a)+\frac{h^2}{2!}f''(a)+\cdots+\frac{h^{n-1}}{(n-1)!}f^{(n-1)}(a)+\int_0^h\frac{t^{n-1}}{(n-1)!}f^{(n)}(a+h-t)\,dt.$$

This, however, is precisely Taylor's Formula with the third form of remainder.

If the point a about which the function is expanded is $x=0$, the expansion will take the form known as Maclaurin's Formula:

$$f(x)=f(0)+xf'(0)+\frac{x^2}{2!}f''(0)+\cdots+\frac{x^{n-1}}{(n-1)!}f^{(n-1)}(0)+R, \quad (3)$$

$$R=\frac{x^n}{n!}f^{(n)}(\theta x)=\frac{x^n}{(n-1)!}(1-\theta)^{n-1}f^{(n)}(\theta x)=\frac{1}{(n-1)!}\int_0^x t^{n-1}f^{(n)}(x-t)\,dt.$$

32. Both Taylor's Formula and its special case, Maclaurin's, express a function as a polynomial in $h=x-a$, of which all the coefficients except the last are constants while the last is not constant but depends on h both explicitly and through the unknown fraction θ which itself is a function of h. If, however, the nth derivative is continuous, the coefficient $f^{(n)}(a+\theta h)/n!$ must remain finite, and if the form of the derivative is known, it may be possible actually to assign limits between which $f^{(n)}(a+\theta h)/n!$ lies. This is of great importance in making approximate calculations as in Exs. 8 ff. below; for it sets a limit to the value of R for any value of n.

Theorem. There is only one possible expansion of a function into a polynomial in $h=x-a$ of which all the coefficients except the last are constant and the last finite; and hence if such an expansion is found in any manner, it must be Taylor's (or Maclaurin's).

To prove this theorem consider two polynomials of the nth order

$$c_0+c_1h+c_2h^2+\cdots+c_{n-1}h^{n-1}+c_nh^n=C_0+C_1h+C_2h^2+\cdots+C_{n-1}h^{n-1}+C_nh^n,$$

which represent the same function and hence are equal for all values of h from 0 to $b-a$. It follows that the coefficients must be equal. For let h approach 0.

The terms containing h will approach 0 and hence c_0 and C_0 may be made as nearly equal as desired; and as they are constants, they must be equal. Strike them out from the equation and divide by h. The new equation must hold for all values of h from 0 to $b - a$ with the possible exception of 0. Again let $h \doteq 0$ and now it follows that $c_1 = C_1$. And so on, with all the coefficients. The two developments are seen to be identical, and hence identical with Taylor's.

To illustrate the application of the theorem, let it be required to find the expansion of $\tan x$ about 0 when the expansions of $\sin x$ and $\cos x$ about 0 are given.

$$\sin x = x - \tfrac{1}{6}x^3 + \tfrac{1}{120}x^5 + Px^7, \qquad \cos x = 1 - \tfrac{1}{2}x^2 + \tfrac{1}{24}x^4 + Qx^6,$$

where P and Q remain finite in the neighborhood of $x = 0$. In the first place note that $\tan x$ clearly has an expansion; for the function and its derivatives (which are combinations of $\tan x$ and $\sec x$) are finite and continuous until x approaches $\frac{1}{2}\pi$. By division,

$$\begin{array}{r|l}
 & x + \tfrac{1}{3}x^3 + \tfrac{2}{15}x^5 \\
\hline
1 - \tfrac{1}{2}x^2 + \tfrac{1}{24}x^4 + Qx^6 & x - \tfrac{1}{6}x^3 + \tfrac{1}{120}x^5 \;\vdots + Px^7 \\
 & x - \tfrac{1}{2}x^3 + \tfrac{1}{24}x^5 \;\vdots + Qx^7 \\
\hline
 & \tfrac{1}{3}x^3 - \tfrac{1}{30}x^5 \;\vdots + (P - Q)x^7 \\
 & \tfrac{1}{3}x^3 - \tfrac{1}{6}x^5 \;\vdots + \tfrac{1}{72}x^7 + \tfrac{1}{3}Qx^9 \\
\hline
 & \tfrac{2}{15}x^5 \;\vdots \cdots\cdots
\end{array}$$

Hence $\tan x = x + \frac{1}{3}x^3 + \frac{2}{15}x^5 + \dfrac{S'}{\cos x}x^7$, where S is the remainder in the division and is an expression containing P, Q, and powers of x; it must remain finite if P and Q remain finite. The quotient $S/\cos x$ which is the coefficient of x^7 therefore remains finite near $x = 0$, and the expression for $\tan x$ is the Maclaurin expansion up to terms of the sixth order, plus a remainder.

In the case of functions compounded from simple functions of which the expansion is known, this method of obtaining the expansion by algebraic processes upon the known expansions treated as polynomials is generally shorter than to obtain the result by differentiation. The computation may be abridged by omitting the last terms and work such as follows the dotted line in the example above; but if this is done, care must be exercised against carrying the algebraic operations too far or not far enough. In Ex. 5 below, the last terms should be put in and carried far enough to insure that the desired expansion has neither more nor fewer terms than the circumstances warrant.

EXERCISES

1. Assume $R = (b - a)^k P$; show $R = \dfrac{h^n(1-\theta)^{n-k}}{(n-1)!\,k} f^{(n)}(\xi)$.

2. Apply Ex. 5, p. 29, to compare the third form of remainder with the first.

3. Obtain, by differentiation and substitution in (1), three nonvanishing terms:

(α) $\sin^{-1}x$, $a = 0$, (β) $\tanh x$, $a = 0$, (γ) $\tan x$, $a = \frac{1}{4}\pi$,

(δ) $\csc x$, $a = \frac{1}{3}\pi$, (ϵ) $e^{\sin x}$, $a = 0$, (ζ) $\log\sin x$, $a = \frac{1}{2}\pi$.

4. Find the nth derivatives in the following cases and write the expansion:

(α) $\sin x$, $a = 0$, (β) $\sin x$, $a = \frac{1}{2}\pi$, (γ) c^x, $a = 0$,

(δ) c^x, $a = 1$, (ϵ) $\log x$, $a = 1$, (ζ) $(1 + x)^k$, $a = 0$.

5. By algebraic processes find the Maclaurin expansion to the term in x^5:

(α) $\sec x$,	(β) $\tanh x$,	(γ) $-\sqrt{1-x^2}$,
(δ) $e^x \sin x$,	(ϵ) $[\log(1-x)]^2$,	(ζ) $+\sqrt{\cosh x}$,
(η) $e^{\sin x}$,	(θ) $\log\cos x$,	(ι) $\log\sqrt{1+x^2}$.

The expansions needed in this work may be found by differentiation or taken from B. O. Peirce's "Tables." In (γ) and (ζ) apply the binomial theorem of Ex. 4 (ζ). In (η) let $y = \sin x$, expand e^y, and substitute for y the expansion of $\sin x$. In (θ) let $\cos x = 1 - y$. In all cases show that the coefficient of the term in x^6 really remains finite when $x \doteq 0$.

6. If $f(a+h) = c_0 + c_1h + c_2h^2 + \cdots + c_{n-1}h^{n-1} + c_nh^n$, show that in

$$\int_0^h f(a+h)\,dh = c_0h + \frac{c_1}{2}h^2 + \frac{c_2}{3}h^3 + \cdots + \frac{c_{n-1}}{n}h^n + \int_0^h c_nh^n dh$$

the last term may really be put in the form Ph^{n+1} with P finite. Apply Ex. 5, p. 29.

7. Apply Ex. 6 to $\sin^{-1}x = \int_0^x \frac{dx}{\sqrt{1-x^2}}$, etc., to find developments of

(α) $\sin^{-1}x$,	(β) $\tan^{-1}x$,	(γ) $\sinh^{-1}x$,
(δ) $\log\frac{1+x}{1-x}$,	(ϵ) $\int_0^x e^{-x^2}dx$,	(ζ) $\int_0^x \frac{\sin x}{x}dx$.

In all these cases the results may be found if desired to n terms.

8. Show that the remainder in the Maclaurin development of e^x is less than $x^ne^x/n!$; and hence that the error introduced by disregarding the remainder in computing e^x is less than $x^ne^x/n!$. How many terms will suffice to compute e to four decimals? How many for e^5 and for $e^{0.1}$?

9. Show that the error introduced by disregarding the remainder in computing $\log(1+x)$ is not greater than x^n/n if $x > 0$. How many terms are required for the computation of $\log 1\frac{1}{2}$ to four places? of $\log 1.2$? Compute the latter.

10. The hypotenuse of a triangle is 20 and one angle is 31°. Find the sides by expanding $\sin x$ and $\cos x$ about $a = \frac{1}{6}\pi$ as linear functions of $x - \frac{1}{6}\pi$. Examine the term in $(x - \frac{1}{6}\pi)^2$ to find a maximum value to the error introduced by neglecting it.

11. Compute to 6 places: (α) $e^{\frac{1}{5}}$, (β) $\log 1.1$, (γ) $\sin 30'$, (δ) $\cos 30'$. During the computation one place more than the desired number should be carried along in the arithmetic work for safety.

12. Show that the remainder for $\log(1+x)$ is less than $x^n/n(1+x)^n$ if $x < 0$. Compute (α) $\log 0.9$ to 5 places, (β) $\log 0.8$ to 4 places.

13. Show that the remainder for $\tan^{-1}x$ is less than x^n/n where n may always be taken as odd. Compute to 4 places $\tan^{-1}\frac{1}{5}$.

14. The relation $\frac{1}{4}\pi = \tan^{-1}1 = 4\tan^{-1}\frac{1}{5} - \tan^{-1}\frac{1}{239}$ enables $\frac{1}{4}\pi$ to be found easily from the series for $\tan^{-1}x$. Find $\frac{1}{4}\pi$ to 7 places (intermediate work carried to 8 places).

15. *Computation of logarithms.* (α) If $a = \log\frac{10}{9}$, $b = \log\frac{25}{24}$, $c = \log\frac{81}{80}$, then

$$\log 2 = 7a - 2b + 3c, \qquad \log 3 = 11a - 3b + 5c, \qquad \log 5 = 16a - 4b + 7c.$$

Now $a = -\log(1 - \frac{1}{10})$, $b = -\log(1 - \frac{4}{100})$, $c = \log(1 + \frac{1}{80})$ are readily computed and hence log 2, log 3, log 5 may be found. Carry the calculations of a, b, c to 10 places and deduce the logarithms of 2, 3, 5, 10, retaining only 8 places. Compare Peirce's "Tables," p. 109.

(β) Show that the error in the series for $\log \dfrac{1+x}{1-x}$ is less than $\dfrac{2x^n}{n(1-x)^n}$. Compute log 2 corresponding to $x = \frac{1}{3}$ to 4 places, $\log 1\frac{1}{4}$ to 5 places, $\log 1\frac{1}{5}$ to 6 places.

(γ) Show $\log \dfrac{p}{q} = 2\left[\dfrac{p-q}{p+q} + \dfrac{1}{3}\left(\dfrac{p-q}{p+q}\right)^3 + \cdots + \dfrac{1}{2n-1}\left(\dfrac{p-q}{p+q}\right)^{2n-1} + R_{2n+1}\right]$, give an estimate of R_{2n+1}, and compute to 10 figures log 3 and log 7 from log 2 and log 5 of Peirce's "Tables" and from

$$4\log 3 - 4\log 2 - \log 5 = \log\frac{81}{80}, \qquad 4\log 7 - 5\log 2 - \log 3 - 2\log 5 = \log\frac{7^4}{7^4 - 1}.$$

16. Compute Ex. 7 (ϵ) to 4 places for $x = 1$ and to 6 places for $x = \frac{1}{2}$.

17. Compute $\sin^{-1} 0.1$ to seconds and $\sin^{-1}\frac{1}{3}$ to minutes.

18. Show that in the expansion of $(1 + x)^k$ the remainder, as x is $>$ or < 0, is

$$R_n < \left|\frac{k\cdot(k-1)\cdots(k-n)}{1\cdot 2\cdots n}x^n\right| \quad \text{or} \quad R_n < \left|\frac{k\cdot(k-1)\cdots(k-n)}{1\cdot 2\cdots n}\,\frac{x^n}{(1+x)^{n-k}}\right|,\ n > k.$$

Hence compute to 5 figures $\sqrt{103}$, $\sqrt{98}$, $\sqrt[3]{28}$, $\sqrt[5]{250}$, $\sqrt[10]{1000}$.

19. Sometimes the remainder cannot be readily found but the terms of the expansion appear to be diminishing so rapidly that all after a certain point appear negligible. Thus use Peirce's "Tables," Nos. 774–789, to compute to four places (estimated) the values of tan 6°, log cos 10°, csc 3°, sec 2°.

20. Find to within 1% the area under $\cos(x^2)$ and $\sin(x^2)$ from 0 to $\frac{1}{2}\pi$.

21. A unit magnetic pole is placed at a distance L from the center of a magnet of pole strength M and length $2l$, where l/L is small. Find the force on the pole if (α) the pole is in the line of the magnet and if (β) it is in the perpendicular bisector.

Ans. (α) $\dfrac{4Ml}{L^3}(1+\epsilon)$ with ϵ about $2\left(\dfrac{l}{L}\right)^2$, ($\beta$) $\dfrac{2Ml}{L^3}(1-\epsilon)$ with ϵ about $\dfrac{3}{2}\left(\dfrac{l}{L}\right)^2$.

22. The formula for the distance of the horizon is $D = \sqrt{\frac{3}{2}h}$ where D is the distance in miles and h is the altitude of the observer in feet. Prove the formula and show that the error is about $\frac{1}{2}$% for heights up to a few miles. Take the radius of the earth as 3960 miles.

23. Find an approximate formula for the dip of the horizon in minutes below the horizontal if h in feet is the height of the observer.

24. If S is a circular arc and C its chord and c the chord of half the arc, prove $S = \frac{1}{3}(8c - C)(1 + \epsilon)$ where ϵ is about $S^4/7680R^4$ if R is the radius.

25. If two quantities differ from each other by a small fraction ϵ of their value, show that their geometric mean will differ from their arithmetic mean by about $\frac{1}{8}\epsilon^2$ of its value.

26. The algebraic method may be applied to finding expansions of some functions which become infinite. (Thus if the series for $\cos x$ and $\sin x$ be divided to find $\cot x$, the initial term is $1/x$ and becomes infinite at $x = 0$ just as $\cot x$ does.

Such expansions are not Maclaurin developments but are analogous to them. The function $x \cot x$ would, however, have a Maclaurin development and the expansion found for $\cot x$ is this development divided by x.) Find the developments about $x = 0$ to terms in x^4 for

(α) $\cot x$, (β) $\cot^2 x$, (γ) $\csc x$, (δ) $\csc^3 x$,
(ϵ) $\cot x \csc x$, (ζ) $1/(\tan^{-1} x)^2$, (η) $(\sin x - \tan x)^{-1}$.

27. Obtain the expansions:

(α) $\log \sin x = \log x - \frac{1}{6} x^2 - \frac{1}{180} x^4 + R$, ($\beta$) $\log \tan x = \log x + \frac{1}{3} x^2 + \frac{7}{90} x^4 + \cdots$,
(γ) likewise for $\log \operatorname{vers} x$.

33. Indeterminate forms, infinitesimals, infinites. If two functions $f(x)$ and $\phi(x)$ are defined for $x = a$ and if $\phi(a) \neq 0$, the quotient f/ϕ is defined for $x = a$. But if $\phi(a) = 0$, the quotient f/ϕ is not defined for a. If in this case f and ϕ are defined and continuous in the neighborhood of a and $f(a) \neq 0$, the quotient will become infinite as $x \doteq a$; whereas if $f(a) = 0$, the behavior of the quotient f/ϕ is not immediately apparent but gives rise to the indeterminate form $0/0$. In like manner if f and ϕ become infinite at a, the quotient f/ϕ is not defined, as neither its numerator nor its denominator is defined; thus arises the indeterminate form ∞/∞. The question of determining or evaluating an indeterminate form is merely the question of finding out whether the quotient f/ϕ approaches a limit (and if so, what limit) or becomes positively or negatively infinite when x approaches a.

THEOREM. *L'Hospital's Rule.* If the functions $f(x)$ and $\phi(x)$, which give rise to the indeterminate form $0/0$ or ∞/∞ when $x \doteq a$, are continuous and differentiable in the interval $a < x \leqq b$ and if b can be taken so near to a that $\phi'(x)$ does not vanish in the interval and if the quotient f'/ϕ' of the derivatives approaches a limit or becomes positively or negatively infinite as $x \doteq a$, then the quotient f/ϕ will approach that limit or become positively or negatively infinite as the case may be. Hence *an indeterminate form* $0/0$ or ∞/∞ *may be replaced by the quotient of the derivatives of numerator and denominator.*

CASE I. $f(a) = \phi(a) = 0$. The proof follows from Cauchy's Formula, Ex. 6, p. 49.

For
$$\frac{f(x)}{\phi(x)} = \frac{f(x) - f(a)}{\phi(x) - \phi(a)} = \frac{f'(\xi)}{\phi'(\xi)}, \qquad a < \xi < x.$$

Now if $x \doteq a$, so must ξ, which lies between x and a. Hence if the quotient on the right approaches a limit or becomes positively or negatively infinite, the same is true of that on the left. The necessity of inserting the restrictions that f and ϕ shall be continuous and differentiable and that ϕ' shall not have a root indefinitely near to a is apparent from the fact that Cauchy's Formula is proved only for functions that satisfy these conditions. If the derived form f'/ϕ' should also be indeterminate, the rule could again be applied and the quotient f''/ϕ'' would replace f'/ϕ' with the understanding that proper restrictions were satisfied by f', ϕ', and ϕ''.

CASE II. $f(a) = \phi(a) = \infty$. Apply Cauchy's Formula as follows:

$$\frac{f(x)-f(b)}{\phi(x)-\phi(b)} = \frac{f(x)}{\phi(x)}\,\frac{1-f(b)/f(x)}{1-\phi(b)/\phi(x)} = \frac{f'(\xi)}{\phi'(\xi)}, \qquad \begin{array}{l} a<x<b, \\ x<\xi<b, \end{array}$$

where the middle expression is merely a different way of writing the first. Now suppose that $f'(x)/\phi'(x)$ approaches a limit when $x \doteq a$. It must then be possible to take b so near to a that $f'(\xi)/\phi'(\xi)$ differs from that limit by as little as desired, no matter what value ξ may have between a and b. Now as f and ϕ become infinite when $x \doteq a$, it is possible to take x so near to a that $f(b)/f(x)$ and $\phi(b)/\phi(x)$ are as near zero as desired. The second equation above then shows that $f(x)/\phi(x)$, multiplied by a quantity which differs from 1 by as little as desired, is equal to a quantity $f'(\xi)/\phi'(\xi)$ which differs from the limit of $f'(x)/\phi'(x)$ as $x \doteq a$ by as little as desired. Hence f/ϕ must approach the same limit as f'/ϕ'. Similar reasoning would apply to the supposition that f'/ϕ' became positively or negatively infinite, and the theorem is proved. It may be noted that, by Theorem 16 of § 27, the form f'/ϕ' is sure to be indeterminate. The advantage of being able to differentiate therefore lies wholly in the possibility that the new form be more amenable to algebraic transformation than the old.

The other indeterminate forms $0 \cdot \infty$, 0^0, 1^∞, ∞^0, $\infty - \infty$ may be reduced to the foregoing by various devices which may be indicated as follows:

$$0 \cdot \infty = \frac{0}{\frac{1}{\infty}} = \frac{\infty}{\frac{1}{0}}, \quad 0^0 = e^{\log 0^0} = e^{0 \log 0} = e^{0 \cdot \infty}, \quad \cdots, \quad \infty - \infty = \log e^{\infty - \infty} = \log \frac{e^\infty}{e^\infty}.$$

The case where the variable becomes infinite instead of approaching a finite value a is covered in Ex. 1 below. The theory is therefore completed.

Two methods which frequently may be used to shorten the work of evaluating an indeterminate form are *the method of E-functions* and *the application of Taylor's Formula*. By definition *an E-function for the point $x = a$ is any continuous function which approaches a finite limit other than 0 when $x \doteq a$.* Suppose then that $f(x)$ or $\phi(x)$ or both may be written as the products $E_1 f_1$ and $E_2 \phi_1$. Then the method of treating indeterminate forms need be applied only to f_1/ϕ_1 and the result multiplied by $\lim E_1/E_2$. For example,

$$\lim_{x \doteq a} \frac{x^3 - a^3}{\sin(x-a)} = \lim_{x \doteq a}(x^2 + ax + a^2) \lim_{x \doteq a} \frac{x-a}{\sin(x-a)} = 3a^2 \lim_{x \doteq a} \frac{x-a}{\sin(x-a)} = 3a^2.$$

Again, suppose that in the form 0/0 both numerator and denominator may be developed about $x = a$ by Taylor's Formula. The evaluation is immediate. Thus

$$\frac{\tan x - \sin x}{x^2 \log(1+x)} = \frac{(x + \frac{1}{3}x^3 + Px^5) - (x - \frac{1}{6}x^3 + Qx^5)}{x^2(x - \frac{1}{2}x^2 + Rx^3)} = \frac{\frac{1}{2} + (P-Q)x^2}{1 - \frac{1}{2}x + Rx^2};$$

and now if $x \doteq 0$, the limit is at once shown to be simply $\frac{1}{2}$.

When the functions become infinite at $x = a$, the conditions requisite for Taylor's Formula are not present and there is no Taylor expansion. Nevertheless an expansion may sometimes be obtained by the algebraic method (§ 32) and may frequently be used to advantage. To illustrate, let it be required to evaluate $\cot x - 1/x$ which is of the form $\infty - \infty$ when $x \doteq 0$. Here

$$\cot x = \frac{\cos x}{\sin x} = \frac{1 + \frac{1}{2}x^2 + Px^4}{x - \frac{1}{6}x^3 + Qx^5} = \frac{1}{x}\,\frac{1 - \frac{1}{2}x^2 + Px^4}{1 - \frac{1}{6}x^2 + Qx^4} = \frac{1}{x}\left(1 - \frac{1}{3}x^2 + Sx^4\right),$$

where S remains finite when $x \doteq 0$. If this value be substituted for $\cot x$, then

$$\lim_{x \doteq 0}\left(\cot x - \frac{1}{x}\right) = \lim_{x \doteq 0}\left(\frac{1}{x} - \frac{1}{3}x + Sx^3 - \frac{1}{x}\right) = \lim_{x \doteq 0}\left(-\frac{1}{3}x + Sx^3\right) = 0.$$

34. *An infinitesimal is a variable which is ultimately to approach the limit zero; an infinite is a variable which is to become either positively or negatively infinite.* Thus the increments Δy and Δx are finite quantities, but when they are to serve in the definition of a derivative they must ultimately approach zero and hence may be called infinitesimals. The form $0/0$ represents the quotient of two infinitesimals; * the form ∞/∞, the quotient of two infinites; and $0 \cdot \infty$, the product of an infinitesimal by an infinite. If any infinitesimal α is chosen as the *primary infinitesimal*, a second infinitesimal β is said to be *of the same order* as α if the limit of the quotient β/α exists and is not zero when $\alpha \doteq 0$; whereas if the quotient β/α becomes zero, β is said to be an infinitesimal *of higher order* than α, but *of lower order* if the quotient becomes infinite. If in particular the limit β/α^n exists and is not zero when $\alpha \doteq 0$, then β is said to be *of the nth order relative to* α. The determination of the order of one infinitesimal relative to another is therefore essentially a problem in indeterminate forms. Similar definitions may be given in regard to infinites.

THEOREM. If the quotient β/α of two infinitesimals approaches a limit or becomes infinite when $\alpha \doteq 0$, the quotient β'/α' of two infinitesimals which differ respectively from β and α by infinitesimals of higher order will approach the same limit or become infinite.

THEOREM. *Duhamel's Theorem.* If the sum $\Sigma\alpha_i = \alpha_1 + \alpha_2 + \cdots + \alpha_n$ of n positive infinitesimals approaches a limit when their number n becomes infinite, the sum $\Sigma\beta_i = \beta_1 + \beta_2 + \cdots + \beta_n$, where each β_i differs uniformly from the corresponding α_i by an infinitesimal of higher order, will approach the same limit.

As $\alpha' - \alpha$ is of higher order than α and $\beta' - \beta$ of higher order than β,

$$\lim \frac{\alpha' - \alpha}{\alpha} = 0, \quad \lim \frac{\beta' - \beta}{\beta} = 0 \quad \text{or} \quad \frac{\alpha'}{\alpha} = 1 + \eta, \quad \frac{\beta'}{\beta} = 1 + \zeta,$$

where η and ζ are infinitesimals. Now $\alpha' = \alpha(1 + \eta)$ and $\beta' = \beta(1 + \zeta)$. Hence

$$\frac{\beta'}{\alpha'} = \frac{\beta}{\alpha}\frac{1 + \zeta}{1 + \eta} \quad \text{and} \quad \lim \frac{\beta'}{\alpha'} = \frac{\beta}{\alpha},$$

provided β/α approaches a limit; whereas if β/α becomes infinite, so will β'/α'. In a more complex fraction such as $(\beta - \gamma)/\alpha$ it is *not* permissible to replace β

* It cannot be emphasized too strongly that in the symbol $0/0$ the 0's are merely symbolic for a mode of variation just as ∞ is; they are not actual 0's and some other notation would be far preferable, likewise for $0 \cdot \infty$, 0^0, etc.

and γ individually by infinitesimals of higher order; for $\beta - \gamma$ may itself be of higher order than β or γ. Thus $\tan x - \sin x$ is an infinitesimal of the third order relative to x although $\tan x$ and $\sin x$ are only of the first order. To replace $\tan x$ and $\sin x$ by infinitesimals which differ from them by those of the second order or even of the third order would generally alter the limit of the ratio of $\tan x - \sin x$ to x^3 when $x \doteq 0$.

To prove Duhamel's Theorem the β's may be written in the form

$$\beta_i = \alpha_i(1 + \eta_i), \qquad i = 1, 2, \cdots, n, \qquad |\eta_i| < \epsilon,$$

where the η's are infinitesimals and where all the η's simultaneously may be made less than the assigned ϵ owing to the uniformity required in the theorem. Then

$$|(\beta_1 + \beta_2 + \cdots + \beta_n) - (\alpha_1 + \alpha_2 + \cdots + \alpha_n)| = |\eta_1\alpha_1 + \eta_2\alpha_2 + \cdots + \eta_n\alpha_n| < \epsilon\Sigma\alpha.$$

Hence the sum of the β's may be made to differ from the sum of the α's by less than $\epsilon\Sigma\alpha$, a quantity as small as desired, and as $\Sigma\alpha$ approaches a limit by hypothesis, so $\Sigma\beta$ must approach the same limit. The theorem may clearly be extended to the case where the α's are not all positive provided the sum $\Sigma|\alpha_i|$ of the absolute values of the α's approaches a limit.

35. If $y = f(x)$, the *differential* of y is defined as

$$dy = f'(x)\,\Delta x, \qquad \text{and hence} \qquad dx = 1 \cdot \Delta x. \tag{4}$$

From this definition of dy and dx it appears that $dy/dx = f'(x)$, where the quotient dy/dx is the quotient of two finite quantities of which dx may be assigned at pleasure. This is true if x is the independent variable. If x and y are both expressed in terms of t,

$$x = x(t), \qquad y = y(t), \qquad dx = D_t x\,dt, \qquad dy = D_t y\,dt;$$

and $$\frac{dy}{dx} = \frac{D_t y}{D_t x} = D_x y, \qquad \text{by virtue of (4), § 2.}$$

From this appears the important theorem: *The quotient dy/dx is the derivative of y with respect to x no matter what the independent variable may be.* It is this theorem which really justifies writing the derivative as a fraction and treating the component differentials according to the rules of ordinary fractions. For higher derivatives this is not so, as may be seen by reference to Ex. 10.

As Δy and Δx are regarded as infinitesimals in defining the derivative, it is natural to regard dy and dx as infinitesimals. The difference $\Delta y - dy$ may be put in the form

$$\Delta y - dy = \left[\frac{f(x + \Delta x) - f(x)}{\Delta x} - f'(x)\right]\Delta x, \tag{5}$$

wherein it appears that, when $\Delta x \doteq 0$, the bracket approaches zero. Hence arises the theorem: *If x is the independent variable and if Δy and dy are regarded as infinitesimals, the difference $\Delta y - dy$ is an infinitesimal of higher order than Δx.* This has an application to the subject of change of variable in a definite integral. For if $x = \phi(t)$, then $dx = \phi'(t)\,dt$, and apparently

$$\int_a^b f(x)\,dx = \int_{t_1}^{t_2} f[\phi(t)]\,\phi'(t)\,dt,$$

where $\phi(t_1) = a$ and $\phi(t_2) = b$, so that t ranges from t_1 to t_2 when x ranges from a to b.

But this substitution is too hasty; for the dx written in the integrand is really Δx, which differs from dx by an infinitesimal of higher order when x is not the independent variable. The true condition may be seen by comparing the two sums

$$\sum f(x_i)\,\Delta x_i, \qquad \sum f[\phi(t_i)]\,\phi'(t_i)\,\Delta t_i, \qquad \Delta t = dt,$$

the limits of which are the two integrals above. Now as Δx differs from $dx = \phi'(t)\,dt$ by an infinitesimal of higher order, so $f(x)\,\Delta x$ will differ from $f[\phi(t)]\,\phi'(t)\,dt$ by an infinitesimal of higher order, and with the proper assumptions as to continuity the difference will be uniform. Hence if the infinitesimals $f(x)\,\Delta x$ be all positive, Duhamel's Theorem may be applied to justify the formula for change of variable. To avoid the restriction to positive infinitesimals it is well to replace Duhamel's Theorem by the new

THEOREM. *Osgood's Theorem.* Let $\alpha_1, \alpha_2, \cdots, \alpha_n$ be n infinitesimals and let α_i differ uniformly by infinitesimals of higher order than Δx from the elements $f(x_i)\,\Delta x_i$ of the integrand of a definite integral $\int_a^b f(x)\,dx$, where f is continuous; then the sum $\Sigma\alpha = \alpha_1 + \alpha_2 + \cdots + \alpha_n$ approaches the value of the definite integral as a limit when the number n becomes infinite.

Let $\alpha_i = f(x_i)\,\Delta x_i + \zeta_i \Delta x_i$, where $|\zeta_i| < \epsilon$ owing to the uniformity demanded.

Then $$\left|\sum \alpha_i - \sum f(x_i)\,\Delta x_i\right| = \left|\sum \zeta_i \Delta x_i\right| < \epsilon \sum \Delta x_i = \epsilon(b-a).$$

But as f is continuous, the definite integral exists and one can make

$$\left|\sum f(x_i)\,\Delta x_i - \int_a^b f(x)\,dx\right| < \epsilon, \quad \text{and hence} \quad \left|\sum \alpha_i - \int_a^b f(x)\,dx\right| < \epsilon(b-a+1).$$

It therefore appears that $\Sigma\alpha_i$ may be made to differ from the integral by as little as desired, and $\Sigma\alpha_i$ must then approach the integral as a limit. Now if this theorem be applied to the case of the change of variable and if it be assumed that $f[\phi(t)]$ and $\phi'(t)$ are continuous, the infinitesimals Δx_i and $dx_i = \phi'(t_i)\,dt_i$ will differ uniformly (compare Theorem 18 of § 27 and the above theorem on $\Delta y - dy$) by an infinitesimal of higher order, and so will the infinitesimals $f(x_i)\,\Delta x_i$ and $f[\phi(t_i)]\,\phi'(t_i)\,dt_i$. Hence the change of variable suggested by the hasty substitution is justified.

EXERCISES

1. Show that l'Hospital's Rule applies to evaluating the indeterminate form $f(x)/\phi(x)$ when x becomes infinite and both f and ϕ either become zero or infinite.

2. Evaluate the following forms by differentiation. Examine the quotients for left-hand and for right-hand approach; sketch the graphs in the neighborhood of the points.

(α) $\lim\limits_{x \doteq 0} \dfrac{a^x - b^x}{x}$, (β) $\lim\limits_{x \doteq \frac{1}{4}\pi} \dfrac{\tan x - 1}{x - \frac{1}{4}\pi}$, (γ) $\lim\limits_{x \doteq 0} x \log x$,

(δ) $\lim\limits_{x = \infty} xe^{-x}$, (ϵ) $\lim\limits_{x \doteq 0} (\cot x)^{\sin x}$, (ζ) $\lim\limits_{x \doteq 1} x^{\frac{1}{1-x}}$.

3. Evaluate the following forms by the method of expansions:

(α) $\lim\limits_{x \doteq 0} \left(\dfrac{1}{x^2} - \cot^2 x\right)$, (β) $\lim\limits_{x \doteq 0} \dfrac{e^x - e^{\tan x}}{x - \tan x}$, (γ) $\lim\limits_{x \doteq 1} \dfrac{\log x}{1 - x}$,

(δ) $\lim\limits_{x \doteq 0} (\operatorname{csch} x - \csc x)$, (ϵ) $\lim\limits_{x \doteq 0} \dfrac{x \sin(\sin x) - \sin^2 x}{x^6}$, (ζ) $\lim\limits_{x \doteq 0} \dfrac{e^x - e^{-x} - 2x}{x - \sin x}$.

4. Evaluate by any method:

(α) $\lim\limits_{x \doteq 0} \dfrac{e^x - e^{-x} + 2\sin x - 4x}{x^5}$, (β) $\lim\limits_{x \doteq 0} \left(\dfrac{\tan x}{x}\right)^{\frac{1}{x^2}}$,

(γ) $\lim\limits_{x \doteq 0} \dfrac{x \cos^3 x - \log(1 + x) - \sin^{-1}\frac{1}{2}x^2}{x^3}$, (δ) $\lim\limits_{x \doteq \frac{1}{2}\pi} \dfrac{\log(x - \frac{1}{2}\pi)}{\tan x}$,

(ϵ) $\lim\limits_{x = \infty} \left[x\left(1 + \dfrac{1}{x}\right)^x - ex^2 \log\left(1 + \dfrac{1}{x}\right)\right]$.

5. Give definitions for order as applied to infinites, noting that higher order would mean becoming infinite to a greater degree just as it means becoming zero to a greater degree for infinitesimals. State and prove the theorem relative to quotients of infinites analogous to that given in the text for infinitesimals. State and prove an analogous theorem for the product of an infinitesimal and infinite.

6. Note that if the quotient of two infinites has the limit 1, the difference of the infinites is an infinite of lower order. Apply this to the proof of the resolution in partial fractions of the quotient $f(x)/F(x)$ of two polynomials in case the roots of the denominator are all real. For if $F(x) = (x - a)^k F_1(x)$, the quotient is an infinite of order k in the neighborhood of $x = a$; but the difference of the quotient and $f(a)/(x - a)^k F_1(a)$ will be of lower integral order — and so on.

7. Show that when $x = +\infty$, the function e^x is an infinite of higher order than x^n no matter how large n. Hence show that if $P(x)$ is any polynomial, $\lim\limits_{x = \infty} P(x)e^{-x} = 0$ when $x = +\infty$.

8. Show that $(\log x)^m$ when x is infinite is a weaker infinite than x^n no matter how large m or how small n, supposed positive, may be. What is the graphical interpretation?

9. If P is a polynomial, show that $\lim\limits_{x \doteq 0} P\left(\dfrac{1}{x}\right)e^{-\frac{1}{x^2}} = 0$. Hence show that the Maclaurin development of $e^{-\frac{1}{x^2}}$ is $f(x) = e^{-\frac{1}{x^2}} = \dfrac{x^n}{n!} f^{(n)}(\theta x)$ if $f(0)$ is defined as 0.

10. The higher differentials are defined as $d^n y = f^{(n)}(x)(dx)^n$ where x is taken as the independent variable. Show that $d^k x = 0$ for $k > 1$ if x is the independent variable. Show that the higher derivatives $D_x^2 y$, $D_x^3 y, \cdots$ are not the quotients d^2y/dx^2, $d^3y/dx^3, \cdots$ if x and y are expressed in terms of a third variable, but that the relations are

$$D_x^2 y = \frac{d^2y dx - d^2x dy}{dx^3}, \qquad D_x^3 y = \frac{dx\,(dx d^3y - dy d^3x) - 3\,d^2x\,(dx d^2y - dy d^2x)}{dx^5}, \quad \cdots.$$

The fact that the quotient d^ny/dx^n, $n > 1$, is not the derivative when x and y are expressed parametrically militates against the usefulness of the higher differentials and emphasizes the advantage of working with derivatives. The notation d^ny/dx^n is, however, used for the derivative. Nevertheless, as indicated in Exs. 16–19, higher differentials may be used if proper care is exercised.

11. Compare the conception of higher differentials with the work of Ex. 5, p. 48.

12. Show that in a circle the difference between an infinitesimal arc and its chord is of the third order relative to either arc or chord.

13. Show that if β is of the nth order with respect to α, and γ is of the first order with respect to α, then β is of the nth order with respect to γ.

14. Show that the order of a product of infinitesimals is equal to the sum of the orders of the infinitesimals when all are referred to the same primary infinitesimal α. Infer that in a product each infinitesimal may be replaced by one which differs from it by an infinitesimal of higher order than it without affecting the order of the product.

15. Let A and B be two points of a unit circle and let the angle AOB subtended at the center be the primary infinitesimal. Let the tangents at A and B meet at T, and OT cut the chord AB in M and the arc AB in C. Find the trigonometric expression for the infinitesimal difference $TC - CM$ and determine its order.

16. Compute $d^2(x \sin x) = (2\cos x - x \sin x)\,dx^2 + (\sin x + x\cos x)\,d^2x$ by taking the differential of the differential. Thus find the second derivative of $x \sin x$ if x is the independent variable and the second derivative with respect to t if $x = 1 + t^2$.

17. Compute the first, second, and third differentials, $d^2x \neq 0$.

(α) $x^2 \cos x$, (β) $\sqrt{1-x}\log(1-x)$, (γ) $xe^{2x}\sin x$.

18. In Ex. 10 take y as the independent variable and hence express $D_x^2 y$, $D_x^3 y$ in terms of $D_y x$, $D_y^2 x$. Cf. Ex. 10, p. 14.

19. Make the changes of variable in Exs. 8, 9, 12, p. 14, by the method of differentials, that is, by replacing the derivatives by the corresponding differential expressions where x is not assumed as independent variable and by replacing these differentials by their values in terms of the new variables where the higher differentials of the new independent variable are set equal to 0.

20. Reconsider some of the exercises at the end of Chap. I, say, 17–19, 22, 23, 27, from the point of view of Osgood's Theorem instead of the Theorem of the Mean.

21. Find the areas of the bounding surfaces of the solids of Ex. 11, p. 18.

22. Assume the law $F = kmm'/r^2$ of attraction between particles. Find the attraction of:

(α) a circular wire of radius a and of mass M on a particle m at a distance r from the center of the wire along a perpendicular to its plane; *Ans.* $kMmr(a^2 + r^2)^{-\frac{3}{2}}$.

(β) a circular disk, etc., as in (α); *Ans.* $2kMma^{-2}(1 - r/\sqrt{r^2 + a^2})$.

(γ) a semicircular wire on a particle at its center; *Ans.* $2kMm/\pi a^2$.

(δ) a finite rod upon a particle not in the line of the rod. The answer should be expressed in terms of the angle the rod subtends at the particle.

(ϵ) two parallel equal rods, forming the opposite sides of a rectangle, on each other.

23. Compare the method of derivatives (§ 7), the method of the Theorem of the Mean (§ 17), and the method of infinitesimals above as applied to obtaining the formulas for (α) area in polar coördinates, (β) mass of a rod of variable density, (γ) pressure on a vertical submerged bulkhead, (δ) attraction of a rod on a particle. Obtain the results by each method and state which method seems preferable for each case.

24. Is the substitution $dx = \phi'(t)\,dt$ in the indefinite integral $\int f(x)\,dx$ to obtain the indefinite integral $\int f[\phi(t)]\,\phi'(t)\,dt$ justifiable immediately?

36. Infinitesimal analysis. To work rapidly in the applications of calculus to problems in geometry and physics and to follow readily the books written on those subjects, it is necessary to have some familiarity with working directly with infinitesimals. It is possible by making use of the Theorem of the Mean and allied theorems to retain in every expression its complete exact value; but if that expression is an infinitesimal which is ultimately to enter into a quotient or a limit of a sum, any infinitesimal which is of higher order than that which is ultimately kept will not influence the result and may be discarded at any stage of the work if the work may thereby be simplified. A few theorems worked through by the infinitesimal method will serve partly to show how the method is used and partly to establish results which may be of use in further work. The theorems which will be chosen are:

1. The increment Δx and the differential dx of a variable differ by an infinitesimal of higher order than either.

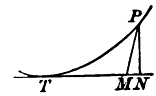
2. If a tangent is drawn to a curve, the perpendicular from the curve to the tangent is of higher order than the distance from the foot of the perpendicular to the point of tangency.

3. An infinitesimal arc differs from its chord by an infinitesimal of higher order relative to the arc.

4. If one angle of a triangle, none of whose angles are infinitesimal, differs infinitesimally from a right angle and if h is the side opposite and if ϕ is another angle of the triangle, then the side opposite ϕ is $h \sin \phi$ except for an infinitesimal of the second order and the adjacent side is $h \cos \phi$ except for an infinitesimal of the first order.

The first of these theorems has been proved in § 35. The second follows from it and from the idea of tangency. For take the x-axis coincident with the tangent or parallel to it. Then the perpendicular is Δy and the distance from its foot to the point of tangency is Δx. The quotient $\Delta y/\Delta x$ approaches 0 as its limit because the tangent is horizontal; and the theorem is proved. *The theorem would remain true if the perpendicular were replaced by a line making a constant angle with the tangent and the distance from the point of tangency to the foot of the perpendicular were replaced by the distance to the foot of the oblique line.* For if $\angle PMN = \theta$,

$$\frac{PM}{TM} = \frac{PN \csc\theta}{TN - PN\cot\theta} = \frac{PN}{TN}\,\frac{\csc\theta}{1 - \dfrac{PN}{TN}\cot\theta},$$

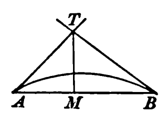


and therefore when P approaches T with θ constant, PM/TM approaches zero and PM is of higher order than TM.

The third theorem follows without difficulty from the assumption or theorem that the arc has a length intermediate between that of the chord and that of the sum of the two tangents at the ends of the chord. Let θ_1 and θ_2 be the angles between the chord and the tangents. Then

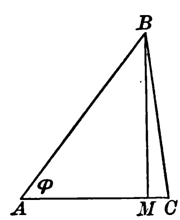
$$\frac{s - AB}{AM + MB} < \frac{AT + TB - AB}{AM + MB} = \frac{AM(\sec\theta_1 - 1) + MB(\sec\theta_2 - 1)}{AM + MB}. \tag{6}$$

Now as AB approaches 0, both $\sec\theta_1 - 1$ and $\sec\theta_2 - 1$ approach 0 and their coefficients remain necessarily finite. Hence the difference between the arc and the chord is an infinitesimal of higher order than the chord. As the arc and chord are therefore of the same order, the difference is of higher order than the arc. This result enables one to replace the arc by its chord and vice versa in discussing infinitesimals of the first order, and for such purposes to consider an infinitesimal arc as straight. In discussing infinitesimals of the second order, this substitution would not be permissible except in view of the further theorem given below in § 37, and even then the substitution will hold only as far as the lengths of arcs are concerned and not in regard to directions.



For the fourth theorem let θ be the angle by which C departs from 90° and with the perpendicular BM as radius strike an arc cutting BC. Then by trigonometry

$$AC = AM + MC = h\cos\phi + BM\tan\theta,$$
$$BC = h\sin\phi + BM(\sec\theta - 1).$$



Now $\tan\theta$ is an infinitesimal of the first order with respect to θ; for its Maclaurin development begins with θ. And $\sec\theta - 1$ is an infinitesimal of the second order; for its development begins with a term in θ^2. The theorem is therefore proved. This theorem is frequently applied to infinitesimal triangles, that is, triangles in which h is to approach 0.

37. As a further discussion of the third theorem it may be recalled that by definition the length of the arc of a curve is the limit of the length of an inscribed polygon, namely,

$$s = \lim_{n=\infty}\left(\sqrt{\Delta x_1^2 + \Delta y_1^2} + \sqrt{\Delta x_2^2 + \Delta y_2^2} + \cdots + \sqrt{\Delta x_n^2 + \Delta y_n^2}\right).$$

Now
$$\sqrt{\Delta x^2+\Delta y^2}-\sqrt{dx^2+dy^2}=\frac{\Delta x^2+\Delta y^2-dx^2-dy^2}{\sqrt{\Delta x^2+\Delta y^2}+\sqrt{dx^2+dy^2}}$$
$$=\frac{(\Delta x-dx)(\Delta x+dx)+(\Delta y-dy)(\Delta y+dy)}{\sqrt{\Delta x^2+\Delta y^2}+\sqrt{dx^2+dy^2}},$$
and
$$\frac{\sqrt{\Delta x^2+\Delta y^2}-\sqrt{dx^2+dy^2}}{\sqrt{\Delta x^2+\Delta y^2}}=\frac{(\Delta x-dx)}{\sqrt{\Delta x^2+\Delta y^2}}\frac{\Delta x+dx}{\sqrt{\Delta x^2+\Delta y^2}+\sqrt{dx^2+dy^2}}$$
$$+\frac{(\Delta y-dy)}{\sqrt{\Delta x^2+\Delta y^2}}\frac{\Delta y+dy}{\sqrt{\Delta x^2+\Delta y^2}+\sqrt{dx^2+dy^2}}.$$

But $\Delta x - dx$ and $\Delta y - dy$ are infinitesimals of higher order than Δx and Δy. Hence the right-hand side must approach zero as its limit and hence $\sqrt{\Delta x^2+\Delta y^2}$ differs from $\sqrt{dx^2+dy^2}$ by an infinitesimal of higher order and may replace it in the sum

$$s=\lim_{n=\infty}\sum\sqrt{\Delta x_i^2+\Delta y_i^2}=\lim_{n=\infty}\sum\sqrt{dx^2+dy^2}=\int_{x_0}^{x_1}\sqrt{1+y'^2}\,dx.$$

The length of the arc measured from a fixed point to a variable point is a function of the upper limit and the differential of arc is

$$ds=d\int_{x_0}^{x}\sqrt{1+y'^2}\,dx=\sqrt{1+y'^2}\,dx=\sqrt{dx^2+dy^2}.$$

To find the order of the difference between the arc and its chord let the origin be taken at the initial point and the x-axis tangent to the curve at that point. The expansion of the arc by Maclaurin's Formula gives

$$s(x)=s(0)+xs'(0)+\tfrac{1}{2}x^2s''(0)+\tfrac{1}{6}x^3s'''(\theta x),$$

where $\quad s(0)=0,\qquad s'(0)=\sqrt{1+y'^2}\,\big|_0=1,\qquad s''(0)=\dfrac{y'y''}{\sqrt{1+y'^2}}\Big|_0=0.$

Owing to the choice of axes, the expansion of the curve reduces to

$$y=f(x)=y(0)+xy'(0)+\tfrac{1}{2}x^2y''(\theta x)=\tfrac{1}{2}x^2y''(\theta x),$$

and hence the chord of the curve is

$$c(x)=\sqrt{x^2+y^2}=x\sqrt{1+\tfrac{1}{4}x^2[y''(\theta x)]^2}=x(1+x^2P),$$

where P is a complicated expression arising in the expansion of the radical by Maclaurin's Formula. The difference

$$s(x)-c(x)=[x+\tfrac{1}{6}x^3s'''(\theta x)]-[x(1+x^2P)]=x^3(\tfrac{1}{6}s'''(\theta x)-P).$$

This is an infinitesimal of at least the third order relative to x. Now as both $s(x)$ and $c(x)$ are of the first order relative to x, it follows that the difference $s(x)-c(x)$ must also be of the third order relative to either $s(x)$ or $c(x)$. Note that the proof assumes that y'' is finite at the point considered. This result, which has been found analytically, follows more simply though perhaps less rigorously from the fact that $\sec\theta_1-1$ and $\sec\theta_2-1$ in (6) are infinitesimals of the second order with θ_1 and θ_2.

38. The theory of *contact of plane curves* may be treated by means of Taylor's Formula and stated in terms of infinitesimals. Let two curves $y=f(x)$ and $y=g(x)$ be tangent at a given point and let the origin be chosen at that point with the x-axis tangent to the curves. The Maclaurin developments are

$$y=f(x)=\frac{1}{2}f''(0)x^2+\cdots+\frac{1}{(n-1)!}x^{n-1}f^{(n-1)}(0)+\frac{1}{n!}x^{(n)}f^{(n)}(0)+\cdots$$

$$y=g(x)=\frac{1}{2}g''(0)x^2+\cdots+\frac{1}{(n-1)!}x^{n-1}g^{(n-1)}(0)+\frac{1}{n!}x^n g^{(n)}(0)+\cdots.$$

If these developments agree up to but not including the term in x^n, the difference between the ordinates of the curves is

$$f(x)-g(x)=\frac{1}{n!}x^n[f^{(n)}(0)-g^{(n)}(0)]+\cdots,\qquad f^{(n)}(0)\neq g^{(n)}(0),$$

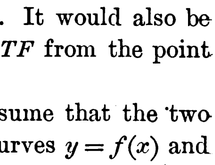
and is an infinitesimal of the nth order with respect to x. The curves are then said to have *contact of order* $n-1$ at their point of tangency. In general when two curves are tangent, the derivatives $f''(0)$ and $g''(0)$ are unequal and the curves have simple contact or *contact of the first order.*

The problem may be stated differently. Let PM be a line which makes a constant angle θ with the x-axis. Then, when P approaches T, if RQ be regarded as straight, the proportion

$$\lim(PR:PQ)=\lim(\sin\angle PQR:\sin\angle PRQ)=\sin\theta:1$$

shows that PR and PQ are of the same order. Clearly also the lines TM and TN are of the same order. Hence if

$$\lim\frac{PR}{(TN)^n}\neq 0,\infty,\quad\text{then}\quad\lim\frac{PQ}{(TM)^n}\neq 0,\infty.$$



Hence if two curves have contact of the $(n-1)$st order, the segment of a line intercepted between the two curves is of the nth order with respect to the distance from the point of tangency to its foot. It would also be of the nth order with respect to the perpendicular TF from the point of tangency to the line.

In view of these results it is not necessary to assume that the two curves have a special relation to the axis. Let two curves $y=f(x)$ and $y=g(x)$ intersect when $x=a$, and assume that the tangents at that point are not parallel to the y-axis. Then

$$y=y_0+(x-a)f'(a)+\cdots+\frac{(x-a)^{n-1}}{(n-1)!}f^{(n-1)}(a)+\frac{(x-a)^n}{n!}f^{(n)}(a)+\cdots$$

$$y=y_0+(x-a)g'(a)+\cdots+\frac{(x-a)^{n-1}}{(n-1)!}f^{(n-1)}(a)+\frac{(x-a)^n}{n!}g^{(n)}(a)+\cdots$$

will be the Taylor developments of the two curves. If the difference of the ordinates for equal values of x is to be an infinitesimal of the nth order with respect to $x-a$ which is the perpendicular from the point of tangency to the ordinate, then the Taylor developments must agree up to but not including the terms in x^n. This is the condition for contact of order $n-1$.

As the difference between the ordinates is

$$f(x)-g(x)=\frac{1}{n!}(x-a)^n[f^{(n)}(a)-g^{(n)}(a)]+\cdots,$$

the difference will change sign or keep its sign when x passes through a according as n is odd or even, because for values sufficiently near to x the higher terms may be neglected. Hence *the curves will cross each other if the order of contact is even,* but *will not cross each other if the order of contact is odd.* If the values of the ordinates are equated to find the points of intersection of the two curves, the result is

$$0=\frac{1}{n!}(x-a)^n\{[f^{(n)}(a)-g^{(n)}(a)]+\cdots\}$$

and shows that $x=a$ is a root of multiplicity n. Hence it is said that two curves have in common as many coincident points as the order of their contact plus one. This fact is usually stated more graphically by saying that *the curves have n consecutive points in common.* It may be remarked that what Taylor's development carried to n terms does, is to give a polynomial which has contact of order $n-1$ with the function that is developed by it.

As a problem on contact consider the determination of the circle which shall have contact of the second order with a curve at a given point (a, y_0). Let

$$y=y_0+(x-a)f'(a)+\tfrac{1}{2}(x-a)^2f''(a)+\cdots$$

be the development of the curve and let $y'=f'(a)=\tan\tau$ be the slope. If the circle is to have contact with the curve, its center must be at some point of the normal. Then if R denotes the assumed radius, the equation of the circle may be written as

$$(x-a)^2+2R\sin\tau\,(x-a)+(y-y_0)^2-2R\cos\tau\,(y-y_0)=0,$$

where it remains to determine R so that the development of the circle will coincide with that of the curve as far as written. Differentiate the equation of the circle.

$$\frac{dy}{dx}=\frac{R\sin\tau+(x-a)}{R\cos\tau-(y-y_0)},\qquad \left(\frac{dy}{dx}\right)_{a,\,y_0}=\tan\tau=f'(a),$$

$$\frac{d^2y}{dx^2}=\frac{[R\cos\tau-(y-y_0)]^2+[R\sin\tau+(x-a)]^2}{[R\cos\tau-(y-y_0)]^3},\qquad \left(\frac{d^2y}{dx^2}\right)_{a,\,y_0}=\frac{1}{R\cos^3\tau},$$

and

$$y=y_0+(x-a)f'(a)+\tfrac{1}{2}(x-a)^2\frac{1}{R\cos^3\tau}+\cdots$$

is the development of the circle. The equation of the coefficients of $(x-a)^2$,

$$\frac{1}{R\cos^3\tau}=f''(a),\quad \text{gives}\quad R=\frac{\sec^3\tau}{f''(a)}=\frac{\{1+[f'(a)]^2\}^{\frac{3}{2}}}{f''(a)}.$$

This is the well known formula for the radius of curvature and shows that the circle of curvature has contact of at least the second order with the curve. The circle is sometimes called the osculating circle instead of the circle of curvature.

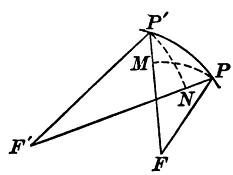
39. Three theorems, one in geometry and two in kinematics, will now be proved to illustrate the direct application of the infinitesimal methods to such problems. The choice will be:

1. The tangent to the ellipse is equally inclined to the focal radii drawn to the point of contact.

2. The displacement of any rigid body in a plane may be regarded at any instant as a rotation through an infinitesimal angle about some point unless the body is moving parallel to itself.

3. The motion of a rigid body in a plane may be regarded as the rolling of one curve upon another.

For the first problem consider a secant PP' which may be converted into a tangent TT' by letting the two points approach until they coincide. Draw the focal radii to P and P' and strike arcs with F and F' as centers. As $F'P+PF=F'P'+P'F=2a$, it follows that $NP=MP'$. Now consider the two triangles $PP'M$ and $P'PN$ nearly right-angled at M and N. The sides PP', PM, PN, $P'M$, $P'N$ are all infinitesimals of the same order and of the same order as the angles at F and F'. By proposition 4 of § 36



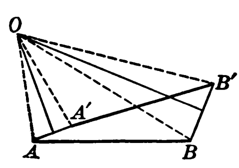
$$MP'=PP'\cos\angle PP'M+e_1,\qquad NP=PP'\cos\angle P'PN+e_2,$$

where e_1 and e_2 are infinitesimals relative to MP' and NP or PP'. Therefore

$$\lim\,[\cos\angle PP'M-\cos\angle P'PN]=\cos\angle TPF'-\cos\angle T'PF=\lim\frac{e_1-e_2}{PP'}=0,$$

and the two angles TPF' and $T'PF$ are proved to be equal as desired.

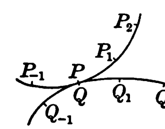
To prove the second theorem note first that if a body is rigid, its position is completely determined when the position AB of any rectilinear segment of the body is known. Let the points A and B of the body be describing curves AA' and BB' so that, in an infinitesimal interval of time, the line AB takes the neighboring position $A'B'$. Erect the perpendicular bisectors of the lines AA' and BB' and let them intersect at O. Then the triangles AOB and $A'OB'$ have the three sides of the one equal to the three sides of the other and are equal, and the second may be obtained from the first by a mere rotation about O through the angle $AOA'=BOB'$. Except for infinitesimals of higher order, the magnitude of the angle is AA'/OA or BB'/OB. Next let the interval of time approach 0 so that A' approaches A and B' approaches B. The perpendicular bisectors will approach



the normals to the arcs AA' and BB' at A and B, and the point O will approach the intersection of those normals.

The theorem may then be stated that: *At any instant of time the motion of a rigid body in a plane may be considered as a rotation through an infinitesimal angle about the intersection of the normals to the paths of any two of its points at that instant; the amount of the rotation will be the distance ds that any point moves divided by the distance of that point from the instantaneous center of rotation; the angular velocity about the instantaneous center will be this amount of rotation divided by the interval of time dt, that is, it will be v/r, where v is the velocity of any point of the body and r is its distance from the instantaneous center of rotation.* It is therefore seen that not only is the desired theorem proved, but numerous other details are found. As has been stated, the point about which the body is rotating at a given instant is called the *instantaneous center* for that instant.

As time goes on, the position of the instantaneous center will generally change. If at each instant of time the position of the center is marked on the moving plane or body, there results a locus which is called the *moving centrode* or *body centrode;* if at each instant the position of the center is also marked on a fixed plane over which the moving plane may be considered to glide, there results another locus which is called the *fixed centrode* or the *space centrode.* From these definitions it follows that at each instant of time the body centrode and the space centrode intersect at the instantaneous center for that instant. Consider a series of positions of the instantaneous center as $P_{-2}P_{-1}PP_1P_2$ marked in space and $Q_{-2}Q_{-1}QQ_1Q_2$ marked in the body. At a given instant two of the points, say P and Q, coincide; an instant later the body will have moved so as to bring Q_1 into coincidence with P_1; at an earlier instant Q_{-1} was coincident with P_{-1}. Now as the motion at the instant when P and Q are together is one of rotation through an infinitesimal angle about that point, the angle between PP_1 and QQ_1 is infinitesimal and the lengths PP_1 and QQ_1 are equal; for it is by the rotation about P and Q that Q_1 is to be brought into coincidence with P_1. Hence it follows 1° that the two centrodes are tangent and 2° that the distances $PP_1 = QQ_1$ which the point of contact moves along the two curves during an infinitesimal interval of time are the same, and this means that the two curves roll on one another without slipping — because the very idea of slipping implies that the point of contact of the two curves should move by different amounts along the two curves, the difference in the amounts being the amount of the slip. The third theorem is therefore proved.



EXERCISES

1. If a finite parallelogram is nearly rectangled, what is the order of infinitesimals neglected by taking the area as the product of the two sides? What if the figure were an isosceles trapezoid? What if it were any rectilinear quadrilateral all of whose angles differ from right angles by infinitesimals of the same order?

2. On a sphere of radius r the area of the zone between the parallels of latitude λ and $\lambda + d\lambda$ is taken as $2\pi r \cos\lambda \cdot r d\lambda$, the perimeter of the base times the slant height. Of what order relative to $d\lambda$ is the infinitesimal neglected? What if the perimeter of the middle latitude were taken so that $2\pi r^2 \cos(\lambda + \frac{1}{2} d\lambda) d\lambda$ were assumed?

3. What is the order of the infinitesimal neglected in taking $4\pi r^2 dr$ as the volume of a hollow sphere of interior radius r and thickness dr? What if the mean radius were taken instead of the interior radius? Would any particular radius be best?

4. Discuss the length of a space curve $y = f(x)$, $z = g(x)$ analytically as the length of the plane curve was discussed in the text.

5. Discuss proposition 2, p. 68, by Maclaurin's Formula and in particular show that if the second derivative is continuous at the point of tangency, the infinitesimal in question is of the second order at least. How about the case of the tractrix

$$y = \frac{a}{2}\log\frac{a - \sqrt{a^2 - x^2}}{a + \sqrt{a^2 - x^2}} + \sqrt{a^2 - x^2},$$

and its tangent at the vertex $x = a$? How about $s(x) - c(x)$ of § 37?

6. Show that if two curves have contact of order $n - 1$, their derivatives will have contact of order $n - 2$. What is the order of contact of the kth derivatives $k < n - 1$?

7. State the conditions for maxima, minima, and points of inflection in the neighborhood of a point where $f^{(n)}(a)$ is the first derivative that does not vanish.

8. Determine the order of contact of these curves at their intersections:

$$(\alpha)\ \begin{matrix}\sqrt{2}(x^2 + y^2 + z) = 3(x + y)\\ 5x^2 - 6xy + 5y^2 = 8,\end{matrix} \qquad (\beta)\ \begin{matrix}r^2 = a^2\cos 2\phi\\ y^2 = \frac{4}{3}a(a - x),\end{matrix} \qquad (\gamma)\ \begin{matrix}x^2 + y^2 = y\\ x^3 + y^3 = xy.\end{matrix}$$

9. Show that at points where the radius of curvature is a maximum or minimum the contact of the osculating circle with the curve must be of at least the third order and must always be of odd order.

10. Let PN be a normal to a curve and $P'N$ a neighboring normal. If O is the center of the osculating circle at P, show with the aid of Ex. 6 that ordinarily the perpendicular from O to $P'N$ is of the second order relative to the arc PP' and that the distance ON is of the first order. Hence interpret the statement: Consecutive normals to a curve meet at the center of the osculating circle.

11. Does the osculating circle cross the curve at the point of osculation? Will the osculating circles at neighboring points of the curve intersect in real points?

12. In the hyperbola the focal radii drawn to any point make equal angles with the tangent. Prove this and state and prove the corresponding theorem for the parabola.

13. Given an infinitesimal arc AB cut at C by the perpendicular bisector of its chord AB. What is the order of the difference $AC - BC$?

14. Of what order is the area of the segment included between an infinitesimal arc and its chord compared with the square on the chord?

15. Two sides AB, AC of a triangle are finite and differ infinitesimally; the angle θ at A is an infinitesimal of the same order and the side BC is either rectilinear or curvilinear. What is the order of the neglected infinitesimal if the area is assumed as $\frac{1}{2}\overline{AB}^2\theta$? What if the assumption is $\frac{1}{2}AB\cdot AC\cdot\theta$?

16. A cycloid is the locus of a fixed point upon a circumference which rolls on a straight line. Show that the tangent and normal to the cycloid pass through the highest and lowest points of the rolling circle at each of its instantaneous positions.

17. Show that the increment of arc Δs in the cycloid differs from $2a\sin\frac{1}{2}\theta d\theta$ by an infinitesimal of higher order and that the increment of area (between two consecutive normals) differs from $3a^2\sin^2\frac{1}{2}\theta d\theta$ by an infinitesimal of higher order. Hence show that the total length and area are $8a$ and $3\pi a^2$. Here a is the radius of the generating circle and θ is the angle subtended at the center by the lowest point and the fixed point which traces the cycloid.

18. Show that the radius of curvature of the cycloid is bisected at the lowest point of the generating circle and hence is $4a\sin\frac{1}{2}\theta$.

19. A triangle ABC is circumscribed about any oval curve. Show that if the side BC is bisected at the point of contact, the area of the triangle will be changed by an infinitesimal of the second order when BC is replaced by a neighboring tangent $B'C'$, but that if BC be not bisected, the change will be of the first order. Hence infer that the minimum triangle circumscribed about an oval will have its three sides bisected at the points of contact.

20. If a string is wrapped about a circle of radius a and then unwound so that its end describes a curve, show that the length of the curve and the area between the curve, the circle, and the string are

$$s = \int_0^\theta a\theta d\theta, \qquad A = \int_0^\theta \tfrac{1}{2}a^2\theta^2 d\theta,$$

where θ is the angle that the unwinding string has turned through.

21. Show that the motion in space of a rigid body one point of which is fixed may be regarded as an instantaneous rotation about some axis through the given point. To do this examine the displacements of a unit sphere surrounding the fixed point as center.

22. Suppose a fluid of variable density $D(x)$ is flowing at a given instant through a tube surrounding the x-axis. Let the velocity of the fluid be a function $v(x)$ of x. Show that during the infinitesimal time δt the diminution of the amount of the fluid which lies between $x = a$ and $x = a + h$ is

$$S[v(a+h)D(a+h)\delta t - v(a)D(a)\delta t],$$

where S is the cross section of the tube. Hence show that $D(x)v(x) = \text{const.}$ is the condition that the flow of the fluid shall not change the density at any point.

23. Consider the curve $y = f(x)$ and three equally spaced ordinates at $x = a - \delta$, $x = a$, $x = a + \delta$. Inscribe a trapezoid by joining the ends of the ordinates at $x = a \pm \delta$ and circumscribe a trapezoid by drawing the tangent at the end of the ordinate at $x = a$ and producing to meet the other ordinates. Show that

$$S_0 = 2\delta f(a), \qquad S = 2\delta\left[f(a) + \frac{\delta^2}{6}f''(a) + \frac{\delta^4}{120}f^{(\text{iv})}(\xi)\right],$$

$$S_1 = 2\delta\left[f(a) + \frac{\delta^2}{2}f''(a) + \frac{\delta^4}{24}f^{(\text{iv})}(\xi_1)\right]$$

are the areas of the circumscribed trapezoid, the curve, the inscribed trapezoid. Hence infer that to compute the area under the curve from the inscribed or circumscribed trapezoids introduces a relative error of the order δ^2, but that to compute from the relation $S = \frac{1}{3}(2S_0 + S_1)$ introduces an error of only the order of δ^4.

24. Let the interval from a to b be divided into an even number $2n$ of equal parts δ and let the $2n+1$ ordinates $y_0, y_1, \cdots, y_{2n}$ at the extremities of the intervals be drawn to the curve $y = f(x)$. Inscribe trapezoids by joining the ends of every other ordinate beginning with y_0, y_2, and going to y_{2n}. Circumscribe trapezoids by drawing tangents at the ends of every other ordinate $y_1, y_3, \cdots, y_{2n-1}$. Compute the area under the curve as

$$S = \int_a^b f(x)\,dx = \frac{b-a}{6}[4(y_1 + y_3 + \cdots + y_{2n-1}) \\ + 2(y_0 + y_2 + \cdots + y_{2n}] - y_0 - y_{2n}] + R$$

by using the work of Ex. 23 and infer that the error R is less than $(b-a)\,\delta^4 f^{(\text{iv})}(\xi)/45$. This method of computation is known as *Simpson's Rule*. It usually gives accuracy sufficient for work to four or even five figures when $\delta = 0.1$ and $b - a = 1$; for $f^{(\text{iv})}(x)$ usually is small.

25. Compute these integrals by Simpson's Rule. Take $2n = 10$ equal intervals. Carry numerical work to six figures except where tables must be used to find $f(x)$:

(α) $\displaystyle\int_1^2 \frac{dx}{x} = \log 2 = 0.69315$, (β) $\displaystyle\int_0^1 \frac{dx}{1+x^2} = \tan^{-1} 1 = \frac{1}{4}\pi = 0.78535$,

(γ) $\displaystyle\int_0^{\frac{1}{2}\pi} \sin x dx = 1.00000$, (δ) $\displaystyle\int_1^2 \log_{10} x dx = 2\log_{10} x - M = 0.16776$,

(ϵ) $\displaystyle\int_0^1 \frac{\log(1+x)}{1+x^2}\,dx = 0.27220$, (ζ) $\displaystyle\int_0^1 \frac{\log(1+x)}{x}\,dx = 0.82247$.

The answers here given are the true values of the integrals to five places.

26. Show that the quadrant of the ellipse $x = a\sin\phi$, $y = b\cos\phi$ is

$$s = a\int_0^{\frac{1}{2}\pi}\sqrt{1 - e^2\sin^2\phi}\,d\phi = \tfrac{1}{2}\pi a\int_0^1 \sqrt{\tfrac{1}{2}(2 - e^2) + \tfrac{1}{2}e^2\cos\pi u}\,du.$$

Compute to four figures by Simpson's Rule with six divisions the quadrants of the ellipses:

(α) $e = \frac{1}{2}\sqrt{3}$, $s = 1.211\,a$, (β) $e = \frac{1}{2}\sqrt{2}$, $s = 1.351\,a$.

27. Expand s in Ex. 26 into a series and discuss the remainder.

$$s = \frac{1}{2}\pi a\left[1 - \left(\frac{1}{2}\right)^2 e^2 - \left(\frac{1\cdot 3}{2\cdot 4}\right)^2\frac{e^4}{3} - \left(\frac{1\cdot 3\cdot 5}{2\cdot 4\cdot 6}\right)^2\frac{e^6}{5} - \cdots\left(\frac{1\cdot 3\cdots(2n-1)}{2\cdot 4\cdots 2n}\right)^2\frac{e^{2n}}{2n-1} - R_n\right]$$

$$R_n < \frac{1}{1-e^2}\left(\frac{1\cdot 3\cdots(2n+1)}{2\cdot 4\cdots(2n+2)}\right)^2\frac{e^{2n+2}}{2n+1}$$ See Ex. 18, p. 60, and Peirce's "Tables," p. 62.

Estimate the number of terms necessary to compute Ex. 26 (β) with an error not greater than 2 in the last place and compare the labor with that of Simpson's Rule.

28. If the eccentricity of an ellipse is $\frac{1}{80}$, find to five decimals the percentage error made in taking $2\pi a$ as the perimeter. *Ans.* 0.00694%

29. If the catenary $y = c\cosh(x/c)$ gives the shape of a wire of length L suspended between two points at the same level and at a distance l nearly equal to L, find the first approximation connecting L, l, and d, where d is the dip of the wire at its lowest point below the level of support.

30. At its middle point the parabolic cable of a suspension bridge 1000 ft. long between the supports sags 50 ft. below the level of the ends. Find the length of the cable correct to inches.

40. Some differential geometry. Suppose that between the increments of a set of variables all of which depend on a single variable t there exists an equation which is true except for infinitesimals of higher order than $\Delta t = dt$, then the equation will be exactly true for the differentials of the variables. Thus if

$$f\Delta x + g\Delta y + h\Delta z + l\Delta t + \cdots + e_1 + e_2 + \cdots = 0$$

is an equation of the sort mentioned and if the coefficients are any functions of the variables and if $e_1, e_2, \cdots$ are infinitesimals of higher order than dt, the limit of

$$f\frac{\Delta x}{\Delta t} + g\frac{\Delta y}{\Delta t} + h\frac{\Delta z}{\Delta t} + l\frac{\Delta t}{\Delta t} + \cdots + \frac{e_1}{\Delta t} + \frac{e_2}{\Delta t} = 0$$

is
$$f\frac{dx}{dt} + g\frac{dy}{dt} + h\frac{dz}{dt} + l = 0,$$

or
$$fdx + gdy + hdz + ldt = 0;$$

and the statement is proved. This result is very useful in writing down various differential formulas of geometry where the approximate relation between the increments is obvious and where the true relation between the differentials can therefore be found.

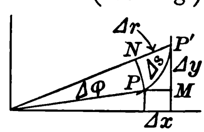
For instance in the case of the differential of arc in rectangular coördinates, if the increment of arc is known to differ from its chord by an infinitesimal of higher order, the Pythagorean theorem shows that the equation

$$\Delta s^2 = \Delta x^2 + \Delta y^2 \quad \text{or} \quad \Delta s^2 = \Delta x^2 + \Delta y^2 + \Delta z^2 \tag{7}$$

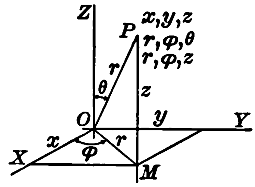
is true except for infinitesimals of higher order; and hence

$$ds^2 = dx^2 + dy^2 \quad \text{or} \quad ds^2 = dx^2 + dy^2 + dz^2. \tag{7'}$$

In the case of plane polar coördinates, the triangle $PP'N$ (see Fig.) has two curvilinear sides PP' and PN and is right-angled at N. The Pythagorean theorem may be applied to a curvilinear triangle, or the triangle may be replaced by the rectilinear triangle $PP'N$ with the angle at N no longer a right angle but nearly so. In either way of looking at the figure, it is easily seen that the equation $\Delta s^2 = \Delta r^2 + r^2\Delta\phi^2$



which the figure suggests differs from a true equation by an infinitesimal of higher order; and hence the inference that in polar coördinates $ds^2 = dr^2 + r^2 d\phi^2$.

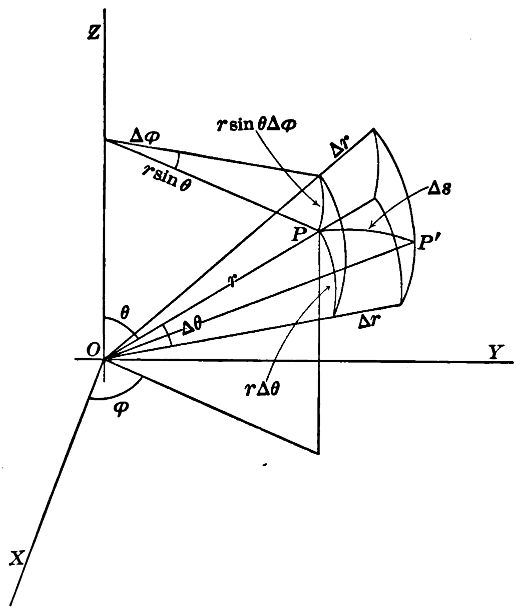


The two most used systems of coördinates other than rectangular in space are the *polar or spherical* and the *cylindrical*. In the first the distance $r = OP$ from the pole or center, the longitude or meridional angle ϕ, and the colatitude or polar angle θ are chosen as coördinates; in the second, ordinary polar coördinates $r = OM$ and ϕ in the xy-plane are combined with the ordinary rectangular z for distance from that plane. The formulas of transformation are

$$z = r\cos\theta, \qquad r = \sqrt{x^2 + y^2 + z^2},$$
$$y = r\sin\theta\sin\phi, \qquad \theta = \cos^{-1}\frac{z}{\sqrt{x^2 + y^2 + z^2}}, \tag{8}$$
$$x = r\sin\theta\cos\phi, \qquad \phi = \tan^{-1}\frac{y}{x},$$

for polar coördinates, and for cylindrical coördinates they are

$$z = z, \quad y = r\sin\phi, \quad x = r\sin\phi, \quad r = \sqrt{x^2 + y^2}, \quad \phi = \tan\frac{y}{x}. \tag{9}$$



Formulas such as that for the differential of arc may be obtained for these new coördinates by mere transformation of (7′) according to the rules for change of variable.

In both these cases, however, the value of ds may be found readily by direct inspection of the figure. The small parallelepiped (figure for polar case) of which Δs is the diagonal has some of its edges and faces curved instead of straight; all the angles, however, are right angles, and as the edges are infinitesimal, the equations certainly suggested as holding except for infinitesimals of higher order are

$$\Delta s^2 = \Delta r^2 + r^2 \sin^2\theta \Delta\phi^2 + r^2\Delta\theta^2 \quad \text{and} \quad \Delta s^2 = \Delta r^2 + r^2\Delta\phi^2 + \Delta z^2 \qquad (10)$$

$$\text{or} \quad ds^2 = dr^2 + r^2 \sin^2\theta d\phi^2 + r^2 d\theta^2 \quad \text{and} \quad ds^2 = dr^2 + r^2 d\phi^2 + dz^2. \qquad (10')$$

To make the proof complete, it would be necessary to show that nothing but infinitesimals of higher order have been neglected and it might actually be easier to transform $\sqrt{dx^2 + dy^2 + dz^2}$ rather than give a rigorous demonstration of this fact. Indeed the infinitesimal method is seldom used rigorously; its great use is to make the facts so clear to the rapid worker that he is willing to take the evidence and omit the proof.

In the plane for rectangular coördinates with rulings parallel to the y-axis and for polar coördinates with rulings issuing from the pole the increments of area differ from

$$dA = ydx \quad \text{and} \quad dA = \tfrac{1}{2} r^2 d\phi \qquad (11)$$

respectively by infinitesimals of higher order, and

$$A = \int_{x_0}^{x_1} ydx \quad \text{and} \quad A = \tfrac{1}{2} \int_{\phi_0}^{\phi_1} r^2 d\phi \qquad (11')$$

are therefore the formulas for the area under a curve and between two ordinates, and for the area between the curve and two radii. If the plane is ruled by lines parallel to both axes or by lines issuing from the pole and by circles concentric with the pole, as is customary for double integration (§§ 131, 134), the increments of area differ respectively by infinitesimals of higher order from

$$dA = dxdy \quad \text{and} \quad dA = rdrd\phi, \qquad (12)$$

and the formulas for the area in the two cases are

$$A = \lim \sum \Delta A = \iint dA = \iint dxdy, \qquad (12')$$

$$A = \lim \sum \Delta A = \iint dA = \iint rdrd\phi,$$

where the double integrals are extended over the area desired.

The elements of volume which are required for triple integration (§§ 133, 134) over a volume in space may readily be written down for the three cases of rectangular, polar, and cylindrical coördinates. In the first case space is supposed to be divided up by planes $x = a$, $y = b$, $z = c$ perpendicular to the axes and spaced at infinitesimal intervals; in the second case the division is made by the spheres $r = a$ concentric with the pole, the planes $\phi = b$ through the polar axis, and the cones $\theta = c$ of revolution about the polar axis; in the third case by the cylinders $r = a$, the planes $\phi = b$, and the planes $z = c$. The infinitesimal volumes into which space is divided then differ from

$$dv = dxdydz, \qquad dv = r^2 \sin\theta dr d\phi d\theta, \qquad dv = r dr d\phi dz \tag{13}$$

respectively by infinitesimals of higher order, and

$$\iiint dxdydz, \qquad \iiint r^2 \sin\theta dr d\phi d\theta, \qquad \iiint r dr d\phi dz \tag{13'}$$

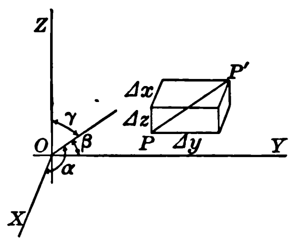
are the formulas for the volumes.

41. The direction of a line in space is represented by the three angles which the line makes with the positive directions of the axes or by the cosines of those angles, the direction cosines of the line. From the definition and figure it appears that

$$l = \cos\alpha = \frac{dx}{ds}, \qquad m = \cos\beta = \frac{dy}{ds}, \qquad n = \cos\gamma = \frac{dz}{dx} \tag{14}$$

are the direction cosines of the tangent to the arc at the point; of the tangent and not of the chord for the reason that the increments are replaced by the differentials. Hence it is seen that for the *direction cosines of the tangent* the proportion

$$l : m : n = dx : dy : dz \tag{14'}$$



holds. The equations of a space curve are

$$x = f(t), \qquad y = g(t), \qquad z = h(t)$$

in terms of a variable parameter t.* At the point (x_0, y_0, z_0) where $t = t_0$ the *equations of the tangent lines* would then be

$$\frac{x - x_0}{(dx)_0} = \frac{y - y_0}{(dy)_0} = \frac{z - z_0}{(dz)_0} \quad \text{or} \quad \frac{x - x_0}{f'(t_0)} = \frac{y - y_0}{g'(t_0)} = \frac{z - z_0}{h'(t_0)}. \tag{15}$$

As the cosine of the angle θ between the two directions given by the direction cosines l, m, n and l', m', n' is

$$\cos\theta = ll' + mm' + nn', \quad \text{so} \quad ll' + mm' + nn' = 0 \tag{16}$$

is the condition for the perpendicularity of the lines. Now if (x, y, z) lies in the plane normal to the curve at x_0, y_0, z_0, the lines determined by the ratios $x - x_0 : y - y_0 : z - z_0$ and $(dx)_0 : (dy)_0 : (dz)_0$ will be perpendicular. Hence the *equation of the normal plane* is

$$(x - x_0)(dx)_0 + (y - y_0)(dy)_0 + (z - z_0)(dz)_0 = 0$$

or

$$f'(t_0)(x - x_0) + g'(t_0)(y - y_0) + h'(t_0)(z - z_0) = 0. \tag{17}$$

* For the sake of generality the parametric form in t is assumed; in a particular case a simplification might be made by taking one of the variables as t and one of the functions f', g', h' would then be 1. Thus in Ex. 8 (ϵ), y should be taken as t.

The *tangent plane* to the curve is not determinate; any plane through the tangent line will be tangent to the curve. If λ be a parameter, the pencil of tangent planes is

$$\frac{x-x_0}{f'(t_0)}+\lambda\frac{y-y_0}{g'(t_0)}-(1+\lambda)\frac{z-z_0}{h'(t_0)}=0.$$

There is one particular tangent plane, called *the osculating plane*, which is of especial importance. Let

$$x-x_0=f'(t_0)\tau+\tfrac{1}{2}f''(t_0)\tau^2+\tfrac{1}{6}f'''(\xi)\tau^3,\quad \tau=t-t_0,\quad \tau_0<\xi<t,$$

with similar expansions for y and z, be the Taylor developments of x, y, z about the point of tangency. When these are substituted in the equation of the plane, the result is

$$\frac{1}{2}\tau^2\left[\frac{f''(t_0)}{f'(t_0)}+\lambda\frac{g''(t_0)}{g'(t_0)}-(1+\lambda)\frac{h''(t_0)}{h'(t_0)}\right]$$
$$+\frac{1}{6}\tau^3\left[\frac{f'''(\xi)}{f'(t_0)}+\lambda\frac{g'''(\eta)}{g'(t_0)}-(1+\lambda)\frac{h'''(\zeta)}{h'(t_0)}\right].$$

This expression is of course proportional to the distance from any point x, y, z of the curve to the tangent plane and is seen to be in general of the second order with respect to τ or ds. It is, however, possible to choose for λ that value which makes the first bracket vanish. The tangent plane thus selected has the property that *the distance of the curve from it in the neighborhood of the point of tangency is of the third order and is called the osculating plane.* The substitution of the value of λ gives

$$\begin{vmatrix} x-x_0 & y-y_0 & z-z_0 \\ f'(t_0) & g'(t_0) & h'(t_0) \\ f''(t_0) & g''(t_0) & h''(t_0) \end{vmatrix}=0 \quad\text{or}\quad \begin{vmatrix} x-x_0 & y-y_0 & z-z_0 \\ (dx)_0 & (dy)_0 & (dz)_0 \\ (d^2x)_0 & (d^2y)_0 & (d^2z)_0 \end{vmatrix}=0 \qquad (18)$$

or
$$(dyd^2z-dzd^2y)_0(x-x_0)+(dzd^2x-dxd^2z)_0(y-y_0)$$
$$+(dxd^2y-dyd^2x)_0(z-z_0)=0$$

as the equation of the osculating plane. In case $f''(t_0)=g''(t_0)=h''(t_0)=0$, this equation of the osculating plane vanishes identically and it is necessary to push the development further (Ex. 11).

42. For the case of plane curves the *curvature* is defined as the rate at which the tangent turns compared with the description of arc, that is, as $d\phi/ds$ if $d\phi$ denotes the differential of the angle through which the tangent turns when the point of tangency advances along the curve by ds. The radius of curvature R is the reciprocal of the curvature, that is, it is $ds/d\phi$. Then

$$d\phi=d\tan^{-1}\frac{dy}{dx},\qquad \frac{d\phi}{ds}=\frac{d\phi}{dx}\frac{dx}{ds}=\frac{y''}{[1+y'^2]^{\frac{3}{2}}},\qquad R=\frac{[1+y'^2]^{\frac{3}{2}}}{y''},\quad (19)$$

where accents denote differentiation with respect to x. For space curves the same definitions are given. If l, m, n and $l + dl$, $m + dm$, $n + dn$ are the direction cosines of two successive tangents,

$$\cos d\phi = l(l + dl) + m(m + dm) + n(n + dn).$$

But $\quad l^2 + m^2 + n^2 = 1 \quad$ and $\quad (l + dl)^2 + (m + dm)^2 + (n + dn)^2 = 1.$

Hence $\quad dl^2 + dm^2 + dn^2 = 2 - 2\cos\phi = (2\sin\frac{1}{2}\phi)^2,$

$$\frac{1}{R^2} = \left(\frac{d\phi}{ds}\right)^2 = \left[\frac{d(2\sin\frac{1}{2}\phi)}{ds}\right]^2 = \frac{dl^2 + dm^2 + dn^2}{ds^2} = l'^2 + m'^2 + n'^2, \quad (19')$$

where accents denote differentiation with respect to s.

The *torsion* of a space curve is defined as the rate of turning of the osculating plane compared with the increase of arc (that is, $d\psi/ds$, where $d\psi$ is the differential angle the normal to the osculating plane turns through), and may clearly be calculated by the same formula as the curvature provided the direction cosines L, M, N of the normal to the plane take the places of the direction cosines l, m, n of the tangent line. Hence the torsion is

$$\frac{1}{\mathsf{R}^2} = \left(\frac{d\psi}{ds}\right)^2 = \frac{dL^2 + dM^2 + dN^2}{ds^2} = L'^2 + M'^2 + N'^2; \qquad (20)$$

and the radius of torsion R is defined as the reciprocal of the torsion, where from the equation of the osculating plane

$$\frac{L}{dyd^2z - dzd^2y} = \frac{M}{dzd^2x - dxd^2z} = \frac{N}{dxd^2y - dyd^2x}$$
$$= \frac{1}{\sqrt{\text{sum of squares}}}. \qquad (20')$$

The actual computation of these quantities is somewhat tedious.

The vectorial discussion of curvature and torsion (§ 77) gives a better insight into the principal directions connected with a space curve. These are the direction of the *tangent*, that of the normal in the osculating plane and directed towards the concave side of the curve and called the *principal normal*, and that of the normal to the osculating plane drawn upon that side which makes the three directions form a right-handed system and called the *binormal*. In the notations there given, combined with those above,

$$\mathbf{r} = x\mathbf{i} + y\mathbf{i} + z\mathbf{k}, \qquad \mathbf{t} = l\mathbf{i} + m\mathbf{j} + n\mathbf{k}, \qquad \mathbf{c} = \lambda\mathbf{i} + \mu\mathbf{j} + \nu\mathbf{k}, \qquad \mathbf{n} = L\mathbf{i} + M\mathbf{j} + N\mathbf{k},$$

where λ, μ, ν are taken as the direction cosines of the principal normal. Now $d\mathbf{t}$ is parallel to $\mathbf{c}$ and $d\mathbf{n}$ is parallel to $-\mathbf{c}$. Hence the results

$$\frac{dl}{\lambda} = \frac{dm}{\mu} = \frac{dn}{\nu} = \frac{ds}{R} \quad \text{and} \quad \frac{dL}{\lambda} = \frac{dM}{\mu} = \frac{dN}{\nu} = -\frac{ds}{\mathsf{R}} \qquad (21)$$

follow from $d\mathbf{c}/ds = \mathbf{C}$ and $d\mathbf{n}/ds = \mathbf{T}$. Now $d\mathbf{c}$ is perpendicular to $\mathbf{c}$ and hence in the plane of $\mathbf{t}$ and $\mathbf{n}$; it may be written as $d\mathbf{c} = (\mathbf{t}\cdot d\mathbf{c})\mathbf{t} + (\mathbf{n}\cdot d\mathbf{c})\mathbf{n}$. But as $\mathbf{t}\cdot\mathbf{c} = \mathbf{n}\cdot\mathbf{c} = 0$, $\mathbf{t}\cdot d\mathbf{c} = -\mathbf{c}\cdot d\mathbf{t}$ and $\mathbf{n}\cdot d\mathbf{c} = -\mathbf{c}\cdot d\mathbf{n}$. Hence

$$d\mathbf{c} = -(\mathbf{c}\cdot d\mathbf{t})\mathbf{t} - (\mathbf{c}\cdot d\mathbf{n})\mathbf{n} = -\mathbf{C}\mathbf{t}ds + \mathbf{T}\mathbf{n}ds = -\frac{\mathbf{t}}{R}ds + \frac{\mathbf{n}}{\mathsf{R}}ds.$$

Hence
$$\frac{d\lambda}{ds} = -\frac{l}{R} + \frac{L}{\mathsf{R}}, \qquad \frac{d\mu}{ds} = -\frac{m}{R} + \frac{M}{\mathsf{R}}, \qquad \frac{d\nu}{ds} = -\frac{n}{R} + \frac{N}{\mathsf{R}}. \tag{22}$$

Formulas (22) are known as *Frenet's Formulas;* they are usually written with $-\mathsf{R}$ in the place of R because a left-handed system of axes is used and the torsion, being an odd function, changes its sign when all the axes are reversed. If accents denote differentiation by s,

$$\text{above formulas, right-handed } \frac{1}{\mathsf{R}} = \frac{\begin{vmatrix} x' & y' & z' \\ x'' & y'' & z'' \\ x''' & y''' & z''' \end{vmatrix}}{x''^2 + y''^2 + z''^2}, \qquad \text{usual formulas, left-handed } \frac{1}{\mathsf{R}} = -\frac{\begin{vmatrix} x' & y' & z' \\ x'' & y'' & z'' \\ x''' & y''' & z''' \end{vmatrix}}{x''^2 + y''^2 + z''^2}. \tag{23}$$

EXERCISES

1. Show that in polar coördinates in the plane, the tangent of the inclination of the curve to the radius vector is $rd\phi/dr$.

2. Verify (10), (10′) by direct transformation of coördinates.

3. Fill in the steps omitted in the text in regard to the proof of (10), (10′) by the method of infinitesimal analysis.

4. A rhumb line on a sphere is a line which cuts all the meridians at a constant angle, say α. Show that for a rhumb line $\sin\theta d\phi = \tan\alpha d\theta$ and $ds = r\sin\alpha d\theta$. Hence find the equation of the line, show that it coils indefinitely around the poles of the sphere, and that its total length is $\pi r \sec\alpha$.

5. Show that the surfaces represented by $F(\phi, \theta) = 0$ and $F(r, \theta) = 0$ in polar coördinates in space are respectively cones and surfaces of revolution about the polar axis. What sort of surface would the equation $F(r, \phi) = 0$ represent?

6. Show accurately that the expression given for the differential of area in polar coördinates in the plane and for the differentials of volume in polar and cylindrical coördinates in space differ from the corresponding increments by infinitesimals of higher order.

7. Show that $\frac{dr}{ds}$, $r\frac{d\theta}{ds}$, $r\sin\theta\frac{d\phi}{ds}$ are the direction cosines of the tangent to a space curve relative to the radius, meridian, and parallel of latitude.

8. Find the tangent line and normal plane of these curves.

(α) $xyz = 1$, $y^2 = x$ at $(1, 1, 1)$, (β) $x = \cos t$, $y = \sin t$, $z = kt$,

(γ) $2ay = x^2$, $6a^2z = x^3$, (δ) $x = t\cos t$, $y = t\sin t$, $z = kt$,

(ϵ) $y = x^2$, $z^2 = 1 - y$, (ζ) $x^2 + y^2 + z^2 = a^2$, $x^2 + y^2 + 2ax = 0$.

9. Find the equation of the osculating plane in the examples of Ex. 8. Note that if x is the independent variable, the equation of the plane is

$$\left(\frac{dy}{dx}\frac{d^2z}{dx^2} - \frac{dz}{dx}\frac{d^2y}{dx^2}\right)_0 (x - x_0) - \left(\frac{d^2z}{dx^2}\right)_0 (y - y_0) + \left(\frac{d^2y}{dx^2}\right)_0 (z - z_0) = 0.$$

10. A space curve passes through the origin, is tangent to the x-axis, and has $z = 0$ as its osculating plane at the origin. Show that

$$x = tf'(0) + \tfrac{1}{2}t^2f''(0) + \cdots, \qquad y = \tfrac{1}{2}t^2g''(0) + \cdots, \qquad z = \tfrac{1}{6}t^3h'''(0) + \cdots$$

will be the form of its Maclaurin development if $t = 0$ gives $x = y = z = 0$.

11. If the 2d, 3d, $\cdots$, $(n - 1)$st derivatives of f, g, h vanish for $t = t_0$ but not all the nth derivatives vanish, show that there is a plane from which the curve departs by an infinitesimal of the $(n + 1)$st order and with which it therefore has contact of order n. Such a plane is called a hyperosculating plane. Find its equation.

12. At what points if any do the curves (β), (γ), (ϵ), (ζ), Ex. 8 have hyperosculating planes and what is the degree of contact in each case?

13. Show that the expression for the radius of curvature is

$$\frac{1}{R} = \sqrt{x''^2 + y''^2 + z''^2} = \frac{[(g'h'' - h'g'')^2 + (h'f'' - f'h'')^2 + (f'g'' - g'f'')^2]^{\frac{1}{2}}}{[f'^2 + g'^2 + h'^2]^{\frac{3}{2}}},$$

where in the first case accents denote differentiation by s, in the second by t.

14. Show that the radius of curvature of a space curve is the radius of curvature of its projection on the osculating plane at the point in question.

15. From Frenet's Formulas show that the successive derivatives of x are

$$x' = l, \qquad x'' = l' = \frac{\lambda}{R}, \qquad x''' = \frac{\lambda'}{R} - \frac{\lambda R'}{R^2} = -\frac{l}{R^2} - \lambda\frac{R'}{R^2} + \frac{L}{R\mathsf{R}},$$

where accents denote differentiation by s. Show that the results for y and z are the same except that m, μ, M or n, ν, N take the places of l, λ, L. Hence infer that for the nth derivatives the results are

$$x^{(n)} = lP_1 + \lambda P_2 + LP_3, \qquad y^{(n)} = mP_1 + \mu P_2 + MP_3, \qquad z^{(n)} = nP_1 + \nu P_2 + NP_3,$$

where P_1, P_2, P_3 are rational functions of R and R and their derivatives by s.

16. Apply the foregoing to the expansion of Ex. 10 to show that

$$x = s - \frac{1}{6R^2}s^3 + \cdots, \qquad y = \frac{s^2}{2R} - \frac{R'}{6R^2}s^3 + \cdots, \qquad z = \frac{s^3}{6R\mathsf{R}} + \cdots,$$

where R and R are the values at the origin where $s = 0$, $l = \mu = N = 1$, and the other six direction cosines m, n, λ, ν, L, M vanish. Find s and write the expansion of the curve of Ex. 8 (γ) in this form.

17. Note that the distance of a point on the curve as expanded in Ex. 16 from the sphere through the origin and with center at the point $(0, R, R'\mathsf{R})$ is

$$\sqrt{x^2 + (y - R)^2 + (z - R'\mathsf{R})^2} - \sqrt{R^2 + R'^2\mathsf{R}^2} = \frac{(x^2 + y^2 - 2Ry + z^2 - 2R'\mathsf{R}z)}{\sqrt{x^2 + (y - R)^2 + (z - R'\mathsf{R})^2} + \sqrt{R^2 + R'^2\mathsf{R}^2}},$$

and consequently is of the fourth order. The curve therefore has contact of the third order with this sphere. Can the equation of this sphere be derived by a limiting process like that of Ex. 18 as applied to the osculating plane?

18. The osculating plane may be regarded as the plane passed through three consecutive points of the curve; in fact it is easily shown that

$$\lim_{\substack{\delta x,\,\delta y,\,\delta z\\ \Delta x,\,\Delta y,\,\Delta z\\ \text{approach } 0}} \begin{vmatrix} x & y & z & 1\\ x_0 & y_0 & z_0 & 1\\ x_0+\delta x & y_0+\delta y & z_0+\delta z & 1\\ x_0+\Delta x & y_0+\Delta y & z_0+\Delta z & 1 \end{vmatrix} = \frac{1}{2}\begin{vmatrix} x-x_0 & y-y_0 & z-z_0\\ (dx)_0 & (dy)_0 & (dz)_0\\ (d^2x)_0 & (d^2y)_0 & (d^2z)_0 \end{vmatrix} = 0.$$

19. Express the radius of torsion in terms of the derivatives of x, y, z by t (Ex. 10, p. 67).

20. Find the direction, curvature, osculating plane, torsion, and osculating sphere (Ex. 17) of the conical helix $x = t\cos t$, $y = t\sin t$, $z = kt$ at $t = 2\pi$.

21. Upon a plane diagram which shows Δs, Δx, Δy, exhibit the lines which represent ds, dx, dy under the different hypotheses that x, y, or s is the independent variable.

CHAPTER IV

PARTIAL DIFFERENTIATION; EXPLICIT FUNCTIONS

43. Functions of two or more variables. The definitions and theorems about functions of more than one independent variable are to a large extent similar to those given in Chap. II for functions of a single variable, and the changes and difficulties which occur are for the most part amply illustrated by the case of two variables. The work in the text will therefore be confined largely to this case and the generalizations to functions involving more than two variables may be left as exercises.

If the value of a variable z is uniquely determined when the values (x, y) of two variables are known, z is said to be a function $z = f(x, y)$ of the two variables. The set of values $[(x, y)]$ or of points $P(x, y)$ of the xy-plane for which z is defined may be any set, but usually consists of all the points in a certain area or region of the plane bounded by a curve which may or may not belong to the region, just as the end points of an interval may or may not belong to it. Thus the function $1/\sqrt{1 - x^2 - y^2}$ is defined for all points within the circle $x^2 + y^2 = 1$, but not for points on the perimeter of the circle. For most purposes it is sufficient to think of the boundary of the region of definition as a polygon whose sides are straight lines or such curves as the geometric intuition naturally suggests.

The first way of representing the function $z = f(x, y)$ geometrically is by *the surface* $z = f(x, y)$, just as $y = f(x)$ was represented by a curve. This method is not available for $u = f(x, y, z)$, a function of three variables, or for functions of a greater number of variables; for space has only three dimensions. A second method of representing the function $z = f(x, y)$ is by its *contour lines* in the xy-plane, that is, the curves $f(x, y) =$ const. are plotted and to each curve is attached the value of the constant. This is the method employed on maps in marking heights above sea level or depths of the ocean below sea level. It is evident that these contour lines are nothing but the projections on the xy-plane of the curves in which the surface $z = f(x, y)$ is cut by the planes $z =$ const. This method is applicable to functions $u = f(x, y, z)$ of three variables. The *contour surfaces* $u =$ const. which are thus obtained are frequently called *equipotential surfaces.* If the function is single valued, the contour lines or surfaces cannot intersect one another.

The *function* $z = f(x, y)$ *is continuous for* (a, b) when either of the following equivalent conditions is satisfied:

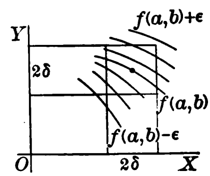
1°. $\lim f(x, y) = f(a, b)$ *or* $\lim f(x, y) = f(\lim x, \lim y)$,

no matter how the variable point $P(x, y)$ *approaches* (a, b).

2°. *If for any assigned* ϵ, *a number* δ *may be found so that*

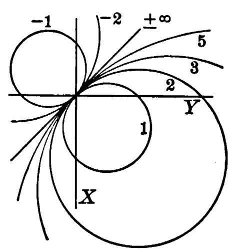
$$|f(x, y) - f(a, b)| < \epsilon \quad \textit{when} \quad |x - a| < \delta,\ |y - b| < \delta.$$

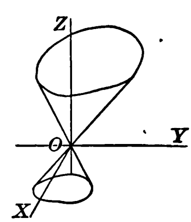
Geometrically this means that if a square with (a, b) as center and with sides of length 2δ parallel to the axes be drawn, the portion of the surface $z = f(x, y)$ above the square will lie between the two planes $z = f(a, b) \pm \epsilon$. Or if contour lines are used, no line $f(x, y) =$ const. where the constant differs from $f(a, b)$ by so much as ϵ will cut into the square. It is clear that in place of a square surrounding (a, b) a circle of radius δ or any other figure which lay within the square might be used.



44. *Continuity examined.* From the definition of continuity just given and from the corresponding definition in § 24, it follows that if $f(x, y)$ is a continuous function of x and y for (a, b), then $f(x, b)$ is a continuous function of x for $x = a$ and $f(a, y)$ is a continuous function of y for $y = b$. That is, if f is continuous in x and y jointly, it is continuous in x and y severally. It might be thought that conversely if $f(x, b)$ is continuous for $x = a$ and $f(a, y)$ for $y = b$, $f(x, y)$ would be continuous in (x, y) for (a, b). That is, if f is continuous in x and y severally, it would be continuous in x and y jointly. A simple example will show that this is *not necessarily true.* Consider the case

$$z = f(x, y) = \frac{x^2 + y^2}{x + y}$$
$$f(0, 0) = 0$$





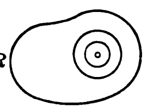
and examine z for continuity at $(0, 0)$. The functions $f(x, 0) = x$, and $f(0, y) = y$ are surely continuous in their respective variables. But the surface $z = f(x, y)$ is a conical surface (except for the points of the z-axis other than the origin) and it is clear that $P(x, y)$ may approach the origin in such a manner that z shall approach any desired value. Moreover, a glance at the contour lines shows that they all enter any circle or square, no matter how small, concentric with the origin. If P approaches the origin along one of these lines, z remains constant and its limiting value is that constant. In fact by approaching the origin along a set of points which jump from one contour line to another, a method of approach may be found such that z approaches no limit whatsoever but oscillates between wide limits or becomes infinite. Clearly the conditions of continuity are not at all fulfilled by z at $(0, 0)$.

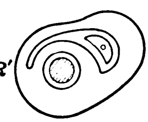
Double limits. There often arise for consideration expressions like

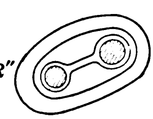
$$\lim_{y \doteq b}\left[\lim_{x \doteq a} f(x, y)\right], \qquad \lim_{x \doteq a}\left[\lim_{y \doteq b} f(x, y)\right], \tag{1}$$

where the limits exist whether x first approaches its limit, and then y its limit, or vice versa, and where the question arises as to whether the two limits thus obtained are equal, that is, whether the order of taking the limits in the double limit may be interchanged. It is clear that if the function $f(x, y)$ is continuous at (a, b), the limits approached by the two expressions will be equal; for the limit of $f(x, y)$ is $f(a, b)$ no matter how (x, y) approaches (a, b). If f is discontinuous at (a, b), it may still happen that the order of the limits in the double limit may be interchanged, as was true in the case above where the value in either order was zero; but this cannot be affirmed in general, and special considerations must be applied to each case when f is discontinuous.

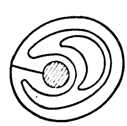
*Varieties of regions.** For both pure mathematics and physics the classification of regions according to their *connectivity* is important. Consider a finite region R bounded by a curve which nowhere cuts itself. (For the present purposes it is not necessary to enter upon the subtleties of the meaning of "curve" (see §§ 127–128); ordinary intuition will suffice.) It is clear that if any closed curve drawn in this region had an unlimited tendency to contract, it could draw together to a point and disappear. On the other hand, if R' be a region like R except that a portion has been removed so that R' is bounded by two curves one within the other, it is clear that some closed curves, namely those which did not encircle the portion removed, could shrink away to a point, whereas other closed curves, namely those which encircled that portion, could at most shrink down into coincidence with the boundary of that portion. Again, if two portions are removed so as to give rise to the region R'', there are circuits around each of the portions which at most can only shrink down to the boundaries of those portions and circuits around both portions which can shrink down to the boundaries and a line joining them. A region like R, where *any* closed curve or circuit may be shrunk away to nothing is called a *simply connected region;* whereas regions in which there are circuits which cannot be shrunk away to nothing are called *multiply connected* regions.

R


R'


R''


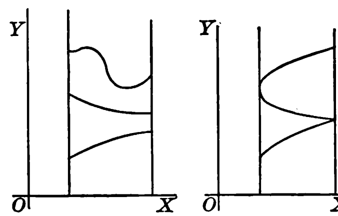
A multiply connected region may be made simply connected by a simple device and convention. For suppose that in R' a line were drawn connecting the two bounding curves and it were agreed that no curve or circuit drawn within R' should cross this line. Then the entire region would be surrounded by a single boundary, part of which would be counted twice. The figure indicates the situation. In like manner if two lines were drawn in R'' connecting both interior boundaries to the exterior or connecting the two interior boundaries together and either of them to the outer boundary, the region would be rendered simply connected. The entire region would have a single boundary of which parts would be counted twice, and any circuit which did not cross the lines could be shrunk away to nothing. The lines



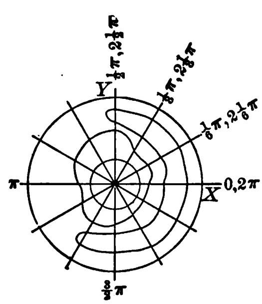
* The discussion from this point to the end of § 45 may be connected with that of §§ 123–126.

thus drawn in the region to make it simply connected are called *cuts*. There is no need that the region be finite; it might extend off indefinitely in some directions like the region between two parallel lines or between the sides of an angle, or like the entire half of the xy-plane for which y is positive. In such cases the cuts may be drawn either to the boundary or off indefinitely in such a way as not to meet the boundary.

45. *Multiple valued functions.* If more than one value of z corresponds to the pair of values (x, y), the function z is multiple valued, and there are some noteworthy differences between multiple valued functions of one variable and of several variables. It was stated (§ 23) that multiple valued functions were divided into branches each of which was single valued. There are two cases to consider when there is one variable, and they are illustrated in the figure. Either there is no value of x in the interval for which the different values of the function are equal and there is consequently a number D which gives the least value of the difference between any two branches, or there is a value of x for which different branches have the same value. Now in the first case, if x changes its value continuously and if $f(x)$ be constrained also to change continuously, there is no possibility of passing from one branch of the function to another; but in the second case such change is possible for, when x passes through the value for which the branches have the same value, the function while constrained to change its value continuously may turn off onto the other branch, although it need not do so.

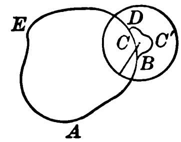


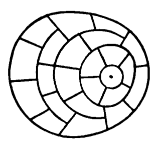
In the case of a function $z = f(x, y)$ of two variables, it is not true that if the values of the function nowhere become equal in or on the boundary of the region over which the function is defined, then it is impossible to pass continuously from one branch to another, and if $P(x, y)$ describes any continuous closed curve or circuit in the region, the value of $f(x, y)$ changing continuously must return to its original value when P has completed the description of the circuit. For suppose the function z be a helicoidal surface $z = a \tan^{-1}(y/x)$, or rather the portion of that surface between two cylindrical surfaces concentric with the axis of the helicoid, as is the case of the surface of the screw of a jack, and the circuit be taken around the inner cylinder. The multiple numbering of the contour lines indicates the fact that the function is multiple valued. Clearly, each time that the circuit is described, the value of z is increased by the amount between the successive branches or leaves of the surface (or decreased by that amount if the circuit is described in the opposite direction). The region here dealt with is not simply connected and the circuit cannot be shrunk to nothing — which is the key to the situation.



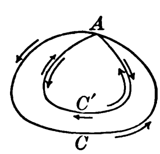
THEOREM. If the difference between the different values of a continuous multiple valued function is never less than a finite number D for any set (x, y) of values of the variables whether in or upon the boundary of the region of definition, then the value $f(x, y)$ of the function, constrained to change continuously, will return to its initial value when the point $P(x, y)$, describing a closed curve which can be shrunk to nothing, completes the circuit and returns to its starting point.

Now owing to the continuity of f throughout the region, it is possible to find a number δ so that $|f(x, y) - f(x', y')| < \epsilon$ when $|x - x'| < \delta$ and $|y - y'| < \delta$ no matter what points of the region (x, y) and (x', y') may be. Hence the values of f at any two points of a small region which lies within any circle of radius $\frac{1}{2}\delta$ cannot differ by so much as the amount D. If, then, the circuit is so small that it may be inclosed within such a circle, there is no possibility of passing from one value of f to another when the circuit is described and f must return to its initial value. Next let there be given any circuit such that the value of f starting from a given value $f(x, y)$ returns to that value when the circuit has been completely described. Suppose that a modification were introduced in the circuit by enlarging or diminishing the inclosed area by a small area lying wholly within a circle of radius $\frac{1}{2}\delta$. Consider the circuit $ABCDEA$ and the modified circuit $ABC'DEA$. As these circuits coincide except for the arcs BCD and $BC'D$, it is only necessary to show that f takes on the same value at D whether D is reached from B by the way of C or by the way of C'. But this is necessarily so for the reason that both arcs are within a circle of radius $\frac{1}{2}\delta$. Then the value of f must still return to its initial value $f(x, y)$ when the modified circuit is described. Now to complete the proof of the theorem, it suffices to note that any circuit which can be shrunk to nothing can be made up by piecing together a number of small circuits as shown in the figure. Then as the change in f around any one of the small circuits is zero, the change must be zero around 2, 3, 4, $\cdots$ adjacent circuits, and thus finally around the complete large circuit.

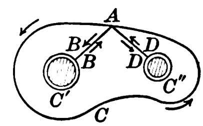


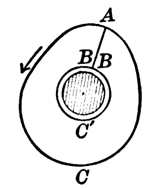


Reducibility of circuits. If a circuit can be shrunk away to nothing, it is said to be *reducible;* if it cannot, it is said to be *irreducible.* In a simply connected region all circuits are reducible; in a multiply connected region there are an infinity of irreducible circuits. Two circuits are said to be *equivalent* or reducible to each other when either can be expanded or shrunk into the other. The change in the value of f on passing around two equivalent circuits from A to A is the same, provided the circuits are described in the same direction. For consider the figure and the equivalent circuits ACA and $AC'A$ described as indicated by the large arrows. It is clear that either may be modified little by little, as indicated in the proof above, until it has been changed into the other. Hence the change in the value of f around the two circuits is the same. Or, as another proof, it may be observed that the combined circuit $ACAC'A$, where the second is described as indicated by the small arrows, may be regarded as a reducible circuit which touches itself at A. Then the change of f around the circuit is zero and f must lose as much on passing from A to A by C' as it gains in passing from A to A by C. Hence on passing from A to A by C' in the direction of the large arrows the gain in f must be the same as on passing by C.



It is now possible to see that *any circuit ABC may be reduced to circuits around the portions cut out of the region combined with lines going to and from A and the boundaries.* The figure shows this; for the circuit $ABC'BADC''DA$ is clearly reducible to the circuit ACA. It must not be forgotten that although the lines AB and BA coincide, the values of the function are not necessarily the same on AB as on BA but differ by the amount of change introduced in f on passing around the irreducible circuit $BC'B$. One of the cases which arises most frequently in practice is that in which the successive branches of $f(x, y)$ differ by a constant amount as in the case $z = \tan^{-1}(y/x)$ where 2π is the difference between successive values of z for the same values of the variables. If now a circuit such as $ABC'BA$ be considered, where it is imagined that the origin lies within $BC'B$, it is clear that the values of z along AB and along BA differ by 2π, and whatever z gains on passing from A to B will be lost on passing from B to A, although the values through which z changes will be different in the two cases by the amount 2π. Hence the circuit $ABC'BA$ gives the same changes for z as the simpler circuit $BC'B$. In other words the result is obtained that *if the different values of a multiple valued function for the same values of the variables differ by a constant independent of the values of the variables, any circuit may be reduced to circuits about the boundaries of the portions removed;* in this case the lines going from the point A to the boundaries and back may be discarded.





EXERCISES

1. Draw the contour lines and sketch the surfaces corresponding to

$$(\alpha)\ z = \frac{x+y}{x-y},\quad z(0, 0) = 0, \qquad (\beta)\ z = \frac{xy}{x+y},\quad z(0, 0) = 0.$$

Note that here and in the text only one of the contour lines passes through the origin although an infinite number have it as a frontier point between two parts of the same contour line. Discuss the double limits $\lim_{x \doteq 0} \lim_{y \doteq 0} z$, $\lim_{y \doteq 0} \lim_{x \doteq 0} z$.

2. Draw the contour lines and sketch the surfaces corresponding to

$$(\alpha)\ z = \frac{x^2 + y^2 - 1}{2y}, \qquad (\beta)\ z = \frac{y^2}{x}, \qquad (\gamma)\ z = \frac{x^2 + 2y^2 - 1}{2x^2 + y^2 - 1}.$$

Examine particularly the behavior of the function in the neighborhood of the apparent points of intersection of different contour lines. Why apparent?

3. State and prove for functions of two independent variables the generalizations of Theorems 6–11 of Chap. II. Note that the theorem on uniformity is proved for two variables by the application of Ex. 9, p. 40, in almost the identical manner as for the case of one variable.

4. Outline definitions and theorems for functions of three variables. In particular indicate the contour surfaces of the functions

$$(\alpha)\ u = \frac{x + y + 2z}{x - y - z}, \qquad (\beta)\ u = \frac{x^2 + y^2 + z^2}{x + y + z}, \qquad (\gamma)\ u = \frac{xy}{z},$$

and discuss the triple limits as x, y, z in different orders approach the origin.

5. Let $z = P(x, y)/Q(x, y)$, where P and Q are polynomials, be a rational function of x and y. Show that if the curves $P = 0$ and $Q = 0$ intersect in any points, all the contour lines of z will converge toward these points; and conversely show that if two different contour lines of z apparently cut in some point, all the contour lines will converge toward that point, P and Q will there vanish, and z will be undefined.

6. If D is the minimum difference between different values of a multiple valued function, as in the text, and if the function returns to its initial value plus $D' \geqq D$ when P describes a circuit, show that it will return to its initial value plus $D' \geqq D$ when P describes the new circuit formed by piecing on to the given circuit a small region which lies within a circle of radius $\frac{1}{2}\delta$.

7. Study the function $z = \tan^{-1}(y/x)$, noting especially the relation between contour lines and the surface. To eliminate the origin at which the function is not defined draw a small circle about the point (0, 0) and observe that the region of the whole xy-plane outside this circle is not simply connected but may be made so by drawing a cut from the circumference off to an infinite distance. Study the variation of the function as P describes various circuits.

8. Study the contour lines and the surfaces due to the functions

$$(\alpha)\ z = \tan^{-1} xy, \qquad (\beta)\ z = \tan^{-1}\frac{1-x^2}{1-y^2}, \qquad (\gamma)\ z = \sin^{-1}(x-y).$$

Cut out the points where the functions are not defined and follow the changes in the functions about such circuits as indicated in the figures of the text. How may the region of definition be made simply connected?

9. Consider the function $z = \tan^{-1}(P/Q)$ where P and Q are polynomials and where the curves $P = 0$ and $Q = 0$ intersect in n points $(a_1, b_1), (a_2, b_2), \cdots, (a_n, b_n)$ but are not tangent (the polynomials have common solutions which are not multiple roots). Show that the value of the function will change by $2k\pi$ if (x, y) describes a circuit which includes k of the points. Illustrate by taking for P/Q the fractions in Ex. 2.

10. Consider regions or volumes in space. Show that there are regions in which some circuits cannot be shrunk away to nothing; also regions in which all circuits may be shrunk away but not all closed surfaces.

46. First partial derivatives. Let $z = f(x, y)$ be a single valued function, or one branch of a multiple valued function, defined for (a, b) and for all points in the neighborhood. If y be given the value b, then z becomes a function $f(x, b)$ of x alone, and if that function has a derivative for $x = a$, that derivative is called the *partial derivative* of $z = f(x, y)$ with respect to x at (a, b). Similarly, if x is held fast and equal to a and if $f(a, y)$ has a derivative when $y = b$, that derivative is called the partial derivative of z with respect to y at (a, b). To obtain these derivatives formally in the case of a given function $f(x, y)$ it is merely necessary to differentiate the function by the ordinary rules, treating y as a constant when finding the derivative with respect to x and x as a constant for the derivative with respect to y. Notations are

$$\frac{\partial f}{\partial x} = \frac{\partial z}{\partial x} = f'_x = f_x = z'_x = D_x f = D_x z = \left(\frac{dz}{dx}\right)_y$$

for the x-derivative with similar ones for the y-derivative. The partial derivatives are the limits of the quotients

$$\lim_{h \doteq 0} \frac{f(a+h, b) - f(a, b)}{h}, \qquad \lim_{k \doteq 0} \frac{f(a, b+k) - f(a, b)}{k}, \tag{2}$$

provided those limits exist. The application of the Theorem of the Mean to the functions $f(x, b)$ and $f(a, y)$ gives

$$\begin{aligned} f(a+h, b) - f(a, b) &= hf'_x(a+\theta_1 h, b), \quad 0 < \theta_1 < 1, \\ f(a, b+k) - f(a, b) &= kf'_y(a, b+\theta_2 k), \quad 0 < \theta_2 < 1, \end{aligned} \tag{3}$$

under the proper but evident restrictions (see § 26).

Two comments may be made. First, some writers denote the partial derivatives by the same symbols dz/dx and dz/dy as if z were a function of only one variable and were differentiated with respect to that variable; and if they desire especially to call attention to the other variables which are held constant, they affix them as subscripts as shown in the last symbol given (p. 93). This notation is particularly prevalent in thermodynamics. As a matter of fact, it would probably be impossible to devise a simple notation for partial derivatives which should be satisfactory for all purposes. The only safe rule to adopt is to use a notation which is sufficiently explicit for the purposes in hand, and at all times to pay careful attention to what the derivative actually means in each case. Second, it should be noted that for points on the boundary of the region of definition of $f(x, y)$ there may be merely right-hand or left-hand partial derivatives or perhaps none at all. For it is necessary that the lines $y = b$ and $x = a$ cut into the region on one side or the other in the neighborhood of (a, b) if there is to be a derivative even one-sided; and at a corner of the boundary it may happen that neither of these lines cuts into the region.

Theorem. If $f(x, y)$ and its derivatives f'_x and f'_y are continuous functions of (x, y) in the neighborhood of (a, b), the increment Δf may be written in any of the three forms

$$\begin{aligned} \Delta f &= f(a+h, b+k) - f(a, b) \\ &= hf'_x(a+\theta_1 h, b) + kf'_y(a+h, b+\theta_2 k) \\ &= hf'_x(a+\theta h, b+\theta k) + kf'_y(a+\theta h, b+\theta k) \\ &= hf'_x(a, b) + kf'_y(a, b) + \zeta_1 h + \zeta_2 k, \end{aligned} \tag{4}$$

where the θ's are proper fractions, the ζ's infinitesimals.

To prove the first form, add and subtract $f(a+h, b)$; then

$$\begin{aligned} \Delta f &= [f(a+h, b) - f(a, b)] + [f(a+h, b+k) - f(a+h, b)] \\ &= hf'_x(a+\theta_1 h, b) + kf'_y(a+h, b+\theta_2 k) \end{aligned}$$

by the application of the Theorem of the Mean for functions of a single variable (§§ 7, 26). The application may be made because the function is continuous and the indicated derivatives exist. Now if the derivatives are also continuous, they may be expressed as

$$f'_x(a+\theta_1 h, b) = f'_x(a, b) + \zeta_1, \qquad f'_y(a+h, b+\theta_2 k) = f'_y(a, b) + \zeta_2$$

where ζ_1, ζ_2 may be made as small as desired by taking h and k sufficiently small. Hence the third form follows from the first. The second form, which is symmetric in the increments h, k, may be obtained by writing $x = a + th$ and $y = b + tk$. Then $f(x, y) = \Phi(t)$. As f is continuous in (x, y), the function Φ is continuous in t and its increment is

$$\Delta\Phi = f(a + \overline{t + \Delta t}\,h,\ b + \overline{t + \Delta t}\,k) - f(a + th,\ b + tk).$$

This may be regarded as the increment of f taken from the point (x, y) with $\Delta t \cdot h$ and $\Delta t \cdot k$ as increments in x and y. Hence $\Delta\Phi$ may be written as

$$\Delta\Phi = \Delta t \cdot hf'_x(a + th,\ b + tk) + \Delta t \cdot kf'_y(a + th,\ b + tk) + \zeta_1\Delta t \cdot h + \zeta_2\Delta t \cdot k.$$

Now if $\Delta\Phi$ be divided by Δt and Δt be allowed to approach zero, it is seen that

$$\lim \frac{\Delta\Phi}{\Delta t} = hf'_x(a + th,\ b + tk) + kf'_y(a + th,\ b + tk) = \frac{d\Phi}{dt}.$$

The Theorem of the Mean may now be applied to Φ to give $\Phi(1) - \Phi(0) = 1 \cdot \Phi'(\theta)$, and hence

$$\begin{aligned}\Phi(1) - \Phi(0) &= f(a + h,\ b + k) - f(a, b)\\ &= \Delta f = hf'_x(a + \theta h,\ b + \theta k) + kf'_y(a + \theta h,\ b + \theta k).\end{aligned}$$

47. The *partial differentials* of f may be defined as

$$\begin{aligned} d_xf &= f'_x\Delta x, \quad \text{so that} \quad dx = \Delta x, \quad \frac{d_xf}{dx} = \frac{\partial f}{\partial x},\\ d_yf &= f'_y\Delta y, \quad \text{so that} \quad dy = \Delta y, \quad \frac{d_yf}{dy} = \frac{\partial f}{\partial y}, \end{aligned} \tag{5}$$

where the indices x and y introduced in d_xf and d_yf indicate that x and y respectively are alone allowed to vary in forming the corresponding partial differentials. The *total differential*

$$df = d_xf + d_yf = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy, \tag{6}$$

which is the sum of the partial differentials, may be defined as that sum; but it is better defined as that part of the increment

$$\Delta f = \frac{\partial f}{\partial x}\Delta x + \frac{\partial f}{\partial y}\Delta y + \zeta_1\Delta x + \zeta_2\Delta y \tag{7}$$

which is obtained by neglecting the terms $\zeta_1\Delta x + \zeta_2\Delta y$, which are of higher order than Δx and Δy. The total differential may therefore be computed by finding the partial derivatives, multiplying them respectively by dx and dy, and adding.

The total differential of $z = f(x, y)$ may be formed for (x_0, y_0) as

$$z - z_0 = \left(\frac{\partial f}{\partial x}\right)_0(x - x_0) + \left(\frac{\partial f}{\partial y}\right)_0(y - y_0), \tag{8}$$

where the values $x - x_0$ and $y - y_0$ are given to the independent differentials dx and dy, and $df = dz$ is written as $z - z_0$. This, however, is the equation of a plane since x and y are independent. The difference $\Delta f - df$ which measures the distance from the plane to the surface along a parallel to the z-axis is of higher order than $\sqrt{\Delta x^2 + \Delta y^2}$; for

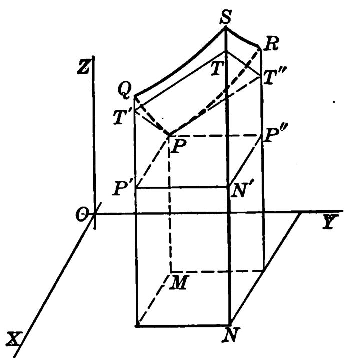
$$\left|\frac{\Delta f - df}{\sqrt{\Delta x^2 + \Delta y^2}}\right| = \left|\frac{\zeta_1 \Delta x + \zeta_2 \Delta y}{\sqrt{\Delta x^2 + \Delta y^2}}\right| < |\zeta_1| + |\zeta_2| \doteq 0.$$

Hence the plane (8) will be defined as the *tangent plane* at (x_0, y_0, z_0) to the surface $z = f(x, y)$. The normal to the plane is

$$\frac{x - x_0}{\left(\frac{\partial f}{\partial x}\right)_0} = \frac{y - y_0}{\left(\frac{\partial f}{\partial y}\right)_0} = \frac{z - z_0}{-1}, \tag{9}$$

which will be defined as the *normal to the surface* at (x_0, y_0, z_0). The tangent plane will cut the planes $y = y_0$ and $x = x_0$ in lines of which the slope is f'_{x_0} and f'_{y_0}. The surface will cut these planes in curves which are tangent to the lines.

In the figure, $PQSR$ is a portion of the surface $z = f(x, y)$ and $PT'TT''$ is a corresponding portion of its tangent plane at $P(x_0, y_0, z_0)$. Now the various values may be read off.



$$PP' = \Delta x, \qquad P'Q = \Delta_x f,$$
$$P'T'/PP' = f'_x, \qquad P'T' = d_x f,$$
$$PP'' = \Delta y, \qquad P''R = \Delta_y f,$$
$$P''T''/PP'' = f'_y, \qquad P''T'' = d_y f,$$
$$P'T' + P''T'' = N'T, \qquad N'S = \Delta f,$$
$$N'T = df = d_x f + d_y f.$$

48. If the variables x and y are expressed as $x = \phi(t)$ and $y = \psi(t)$ so that $f(x, y)$ becomes a function of t, the derivative of f with respect to t is found from the expression for the increment of f.

$$\frac{\Delta f}{\Delta t} = \frac{\partial f}{\partial x}\frac{\Delta x}{\Delta t} + \frac{\partial f}{\partial y}\frac{\Delta y}{\Delta t} + \zeta_1 \frac{\Delta x}{\Delta t} + \zeta_2 \frac{\Delta y}{\Delta t}$$

or

$$\lim_{\Delta t \doteq 0} \frac{\Delta f}{\Delta t} = \frac{df}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt}. \tag{10}$$

The conclusion requires that x and y should have finite derivatives with respect to t. The differential of f as a function of t is

$$df = \frac{df}{dt}dt = \frac{\partial f}{\partial x}\frac{dx}{dt}dt + \frac{\partial f}{\partial y}\frac{dy}{dt}dt = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy \tag{11}$$

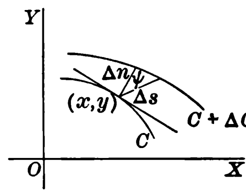
and hence it appears that *the differential has the same form as the total differential.* This result will be generalized later.

As a particular case of (10) suppose that x and y are so related that the point (x, y) moves along a line inclined at an angle τ to the x-axis. If s denote distance along the line, then

$$x = x_0 + s\cos\tau,\quad y = y_0 + s\sin\tau,\quad dx = \cos\tau ds,\quad dy = \sin\tau ds \qquad (12)$$

and
$$\frac{df}{ds} = \frac{\partial f}{\partial x}\frac{dx}{ds} + \frac{\partial f}{\partial y}\frac{dy}{ds} = f'_x \cos\tau + f'_y \sin\tau. \qquad (13)$$

The derivative (13) is called the *directional derivative* of f in the direction of the line. The partial derivatives f'_x, f'_y are the particular directional derivatives along the directions of the x-axis and y-axis. The directional derivative of f in any direction is the rate of increase of f along that direction; if $z = f(x, y)$ be interpreted as a surface, the directional derivative is the slope of the curve in which a plane through the line (12) and perpendicular to the xy-plane cuts the surface. If $f(x, y)$ be represented by its contour lines, the derivative at a point (x, y) in any direction is the limit of the ratio $\Delta f/\Delta s = \Delta C/\Delta s$ of the increase of f, from one contour line to a neighboring one, to the distance between the lines in that direction. It is therefore evident that the derivative along any contour line is zero and that the derivative along the normal to the contour line is greater than in any other direction because the element dn of the normal is less than ds in any other direction. In fact, apart from infinitesimals of higher order,



$$\frac{\Delta n}{\Delta s} = \cos\psi,\quad \frac{\Delta f}{\Delta s} = \frac{\Delta f}{\Delta n}\cos\psi,\quad \frac{df}{ds} = \frac{df}{dn}\cos\psi. \qquad (14)$$

Hence it is seen that *the derivative along any direction may be found by multiplying the derivative along the normal by the cosine of the angle between that direction and the normal.* The derivative along the normal to a contour line is called the *normal derivative* of f and is, of course, a function of (x, y).

49. Next suppose that $u = f(x, y, z, \cdots)$ is a function of any number of variables. The reasoning of the foregoing paragraphs may be repeated without change except for the additional number of variables. The increment of f will take any of the forms

$$\begin{aligned}\Delta f &= f(a + h, b + k, c + l, \cdots) - f(a, b, c, \cdots)\\ &= hf'_x(a + \theta_1 h, b, c, \cdots) + kf'_y(a + h, b + \theta_2 k, c, \cdots)\\ &\quad + lf'_z(a + h, b + k, c + \theta_3 l, \cdots) + \cdots\\ &= [hf'_x + kf'_y + lf'_z + \cdots]_{a+\theta h,\, b+\theta k,\, c+\theta l,\, \cdots}\\ &= hf'_x + kf'_y + lf'_z + \cdots + \zeta_1 h + \zeta_2 k + \zeta_3 l + \cdots,\end{aligned}$$

and the total differential will naturally be defined as

$$df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy + \frac{\partial f}{\partial z} dz + \cdots, \tag{16}$$

and finally if $x, y, z, \cdots$ be functions of t, it follows that

$$\frac{df}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} + \frac{\partial f}{\partial z}\frac{dz}{dt} + \cdots \tag{17}$$

and the differential of f as a function of t is still (16).

If the variables $x, y, z, \cdots$ were expressed in terms of several new variables $r, s, \cdots$, the function f would become a function of those variables. To find the partial derivative of f with respect to one of those variables, say r, the remaining ones, $s, \cdots$, would be held constant and f would for the moment become a function of r alone, and so would x, $y, z, \cdots$. Hence (17) may be applied to obtain the partial derivatives

$$\frac{\partial f}{\partial r} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial r} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial r} + \cdots,$$

and

$$\frac{\partial f}{\partial s} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial s} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial s} + \cdots, \text{ etc.} \tag{18}$$

These are the formulas for *change of variable* analogous to (4) of § 2. If these equations be multiplied by $\Delta r, \Delta s, \cdots$ and added,

$$\frac{\partial f}{\partial r}\Delta r + \frac{\partial f}{\partial s}\Delta s + \cdots = \frac{\partial f}{\partial x}\left(\frac{\partial x}{\partial r}\Delta r + \frac{\partial x}{\partial s}\Delta s + \cdots\right) + \frac{\partial f}{\partial y}\left(\frac{\partial y}{\partial r}\Delta r + \cdots\right) + \cdots,$$

or

$$df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy + \frac{\partial f}{\partial z} dz + \cdots;$$

for when $r, s, \cdots$ are the independent variables, the parentheses above are $dx, dy, dz, \cdots$ and the expression on the left is df.

THEOREM. The expression of the total differential of a function of $x, y, z, \cdots$ as $df = f'_x dx + f'_y dy + f'_z dz + \cdots$ is the same whether x, y, $z, \cdots$ are the independent variables or functions of other independent variables $r, s, \cdots$; it being assumed that all the derivatives which occur, whether of f by $x, y, z, \cdots$ or of $x, y, z, \cdots$ by $r, s, \cdots$, are continuous functions.

By the same reasoning or by virtue of this theorem the rules

$$d(cu) = cdu, \quad d(u + v - w) = du + dv - dw,$$
$$d(uv) = udv + vdu, \quad d\left(\frac{u}{v}\right) = \frac{vdu - udv}{v^2}, \tag{19}$$

of the differential calculus will apply to calculate the total differential of combinations or functions of several variables. If by this means, or any other, there is obtained an expression

$$df = R(r, s, t, \cdots)\,dr + S(r, s, t, \cdots)\,ds + T(r, s, t, \cdots)\,dt + \cdots \quad (20)$$

for the total differential in which $r, s, t, \cdots$ are *independent* variables, the coefficients $R, S, T, \cdots$ are the derivatives

$$R = \frac{\partial f}{\partial r}, \quad S = \frac{\partial f}{\partial s}, \quad T = \frac{\partial f}{\partial t}, \cdots. \quad (21)$$

For in the equation $df = Rdr + Sds + Tdt + \cdots = f'_r dr + f'_s ds + f'_t dt + \cdots$, the variables $r, s, t, \cdots$, being independent, may be assigned increments absolutely at pleasure and if the particular choice $dr = 1$, $ds = dt = \cdots = 0$, be made, it follows that $R = f'_r$; and so on. The single equation (20) is thus equivalent to the equations (21) in number equal to the number of the independent variables.

As an example, consider the case of the function $\tan^{-1}(y/x)$. By the rules (19),

$$d\tan^{-1}\frac{y}{x} = \frac{d(y/x)}{1 + (y/x)^2} = \frac{dy/x - ydx/x^2}{1 + (y/x)^2} = \frac{xdy - ydx}{x^2 + y^2}.$$

Then $$\frac{\partial}{\partial x}\tan^{-1}\frac{y}{x} = -\frac{y}{x^2 + y^2}, \quad \frac{\partial}{\partial y}\tan^{-1}\frac{y}{x} = \frac{x}{x^2 + y^2}, \quad \text{by (20)–(21).}$$

If y and x were expressed as $y = \sinh rst$ and $x = \cosh rst$, then

$$d\tan^{-1}\frac{y}{x} = \frac{xdy - ydx}{x^2 + y^2} = \frac{[stdr + rtds + rsdt][\cosh^2 rst - \sinh^2 rst]}{\cosh^2 rst + \sinh^2 rst}$$

and $$\frac{\partial f}{\partial r} = \frac{st}{\cosh 2\,rst}, \quad \frac{\partial f}{\partial s} = \frac{rt}{\cosh 2\,rst}, \quad \frac{\partial f}{\partial t} = \frac{rs}{\cosh 2\,rst}.$$

EXERCISES

1. Find the partial derivatives f'_x, f'_y or f'_x, f'_y, f'_z of these functions:

(α) $\log(x^2 + y^2)$, (β) $e^x \cos y \sin z$, (γ) $x^2 + 3xy + y^3$,

(δ) $\dfrac{xy}{x + y}$, (ϵ) $\dfrac{e^{xy}}{e^x + e^y}$, ($\zeta$) $\log(\sin x + \sin^2 y + \sin^3 z)$,

(η) $\sin^{-1}\dfrac{y}{x}$, (θ) $\dfrac{z}{x} e^{\frac{y}{x}}$, ($\iota$) $\tanh^{-1}\sqrt{2}\left(\dfrac{xy + yz + zx}{x^2 + y^2 + z^2}\right)^{\frac{1}{2}}$.

2. Apply the definition (2) directly to the following to find the partial derivatives at the indicated points:

(α) $\dfrac{xy}{x + y}$ at (1, 1), (β) $x^2 + 3xy + y^3$ at (0, 0), and (γ) at (1, 1),

(δ) $\dfrac{x - y}{x + y}$ at (0, 0); also try differentiating and substituting (0, 0).

3. Find the partial derivatives and hence the total differential of:

(α) $\dfrac{e^{xy}}{x^2 + y^2}$, ($\beta$) $x \log yz$, (γ) $\sqrt{a^2 - x^2 - y^2}$,

(δ) $e^{-x}\sin y$, (ϵ) $e^{z^2}\sinh xy$, (ζ) $\log\tan\left(x + \dfrac{\pi}{4}y\right)$,

(η) $\left(\dfrac{y}{z}\right)^x$, ($\theta$) $\dfrac{x - y}{x + z}$, (ι) $\log\left(\dfrac{3x}{y^2} + \sqrt{1 + \dfrac{z^2x^2}{y^4}}\right)$.

4. Find the general equations of the tangent plane and normal line to these surfaces and find the equations of the plane and line for the indicated (x_0, y_0):

(α) the helicoid $z = k\tan^{-1}(y/x)$, $(1, 0), (1, -1), (0, 1)$,
(β) the paraboloid $4pz = (x^2 + y^2)$, $(0, p), (2p, 0), (p, -p)$,
(γ) the hemisphere $z = \sqrt{a^2 - x^2 - y^2}$, $(0, -\frac{1}{2}a), (\frac{1}{2}a, \frac{1}{2}a), (\frac{1}{2}\sqrt{3}a, 0)$,
(δ) the cubic $xyz = 1$, $(1, 1, 1), (-\frac{1}{2}, -\frac{1}{2}, 4), (4, \frac{1}{2}, \frac{1}{2})$.

5. Find the derivative with respect to t in these cases by (10):

(α) $f = x^2 + y^2,\ x = a\cos t,\ y = b\sin t$, ($\beta$) $\tan^{-1}\sqrt{\dfrac{y}{x}},\ y = \cosh t,\ x = \sinh t$,

(γ) $\sin^{-1}(x - y),\ x = 3t,\ y = 4t^3$, ($\delta$) $\cos 2xy,\ x = \tan^{-1}t,\ y = \cot^{-1}t$.

6. Find the directional derivative in the direction indicated and obtain its numerical value at the points indicated:

(α) $x^2y,\ \tau = 45°,\ (1, 2)$, ($\beta$) $\sin^2 xy,\ \tau = 60°,\ (\sqrt{3}, -2)$.

7. (α) Determine the maximum value of df/ds from (13) by regarding τ as variable and applying the ordinary rules. Show that the direction that gives the maximum is

$$\tau = \tan^{-1}\frac{f_y'}{f_x'}, \quad \text{and then} \quad \frac{df}{dn} = \sqrt{\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2}.$$

(β) Show that the sum of the squares of the derivatives along any two perpendicular directions is the same and is the square of the normal derivative.

8. Show that $(f_x' + y'f_y')/\sqrt{1 + y'^2}$ and $(f_x'y' - f_y')/\sqrt{1 + y'^2}$ are the derivatives of f along the curve $y = \phi(x)$ and normal to the curve.

9. If df/dn is defined by the work of Ex. 7 (α), prove (14) as a consequence.

10. Apply the formulas for the change of variable to the following cases:

(α) $r = \sqrt{x^2 + y^2},\ \phi = \tan^{-1}\dfrac{y}{x}$. Find $\dfrac{\partial f}{\partial x},\ \dfrac{\partial f}{\partial y},\ \sqrt{\left(\dfrac{\partial f}{\partial x}\right)^2 + \left(\dfrac{\partial f}{\partial y}\right)^2}$.

(β) $x = r\cos\phi,\ y = r\sin\phi$. Find $\dfrac{\partial f}{\partial r},\ \dfrac{\partial f}{\partial \phi},\ \left(\dfrac{\partial f}{\partial r}\right)^2 + \dfrac{1}{r^2}\left(\dfrac{\partial f}{\partial \phi}\right)^2$.

(γ) $x = 2r - 3s + 7,\ y = -r + 8s - 9$. Find $\dfrac{\partial u}{\partial r} = 4x + 2y$ if $u = x^2 - y^2$.

(δ) $\begin{cases} x = x'\cos\alpha - y'\sin\alpha, \\ y = x'\sin\alpha + y'\cos\alpha. \end{cases}$ Show $\left(\dfrac{\partial f}{\partial x}\right)^2 + \left(\dfrac{\partial f}{\partial y}\right)^2 = \left(\dfrac{\partial f}{\partial x'}\right)^2 + \left(\dfrac{\partial f}{\partial y'}\right)^2$.

(ϵ) Prove $\dfrac{\partial f}{\partial x} + \dfrac{\partial f}{\partial y} = 0$ if $f(u, v) = f(x - y, y - x)$.

(ζ) Let $x = ax' + by' + cz',\ y = a'x' + b'y' + c'z',\ z = a''x' + b''y' + c''z'$, where $a, b, c, a', b', c', a'', b'', c''$ are the direction cosines of new rectangular axes with respect to the old. This transformation is called an *orthogonal transformation.* Show

$$\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2 + \left(\frac{\partial f}{\partial z}\right)^2 = \left(\frac{\partial f}{\partial x'}\right)^2 + \left(\frac{\partial f}{\partial y'}\right)^2 + \left(\frac{\partial f}{\partial z'}\right)^2 = \left(\frac{df}{dn}\right)^2.$$

11. Define directional derivative in space; also normal derivative and establish (14) for this case. Find the normal derivative of $f = xyz$ at $(1, 2, 3)$.

12. Find the total differential and hence the partial derivatives in Exs. 1, 3, and

(α) $\log(x^2 + y^2 + z^2)$, (β) y/x, (γ) $x^2ye^{xy^2}$, (δ) $xyz\log xyz$,

(ϵ) $u = x^2 - y^2$, $x = r\cos st$, $y = s\sin rt$. Find $\partial u/\partial r$, $\partial u/\partial s$, $\partial u/\partial t$.

(ζ) $u = y/x$, $x = r\cos\phi\sin\theta$, $y = r\sin\phi\sin\theta$. Find u'_r, u'_ϕ, u'_θ.

(η) $u = e^{xy}$, $x = \log\sqrt{r^2 + s^2}$, $y = \tan^{-1}(s/r)$. Find u'_r, u'_s.

13. If $\frac{\partial f}{\partial x} = \frac{\partial g}{\partial y}$ and $\frac{\partial f}{\partial y} = -\frac{\partial g}{\partial x}$, show $\frac{\partial f}{\partial r} = \frac{1}{r}\frac{\partial g}{\partial \phi}$ and $\frac{1}{r}\frac{\partial f}{\partial \phi} = -\frac{\partial g}{\partial r}$ if r, ϕ are polar coördinates and f, g are any two functions.

14. If $p(x, y, z, t)$ is the pressure in a fluid, or $\rho(x, y, z, t)$ is the density, depending on the position in the fluid and on the time, and if u, v, w are the velocities of the particles of the fluid along the axes,

$$\frac{dp}{dt} = u\frac{\partial p}{\partial x} + v\frac{\partial p}{\partial y} + w\frac{\partial p}{\partial z} + \frac{\partial p}{\partial t} \quad \text{and} \quad \frac{d\rho}{dt} = u\frac{\partial \rho}{\partial x} + v\frac{\partial \rho}{\partial y} + w\frac{\partial \rho}{\partial z} + \frac{\partial \rho}{dt}.$$

Explain the meaning of each derivative and prove the formula.

15. If $z = xy$, interpret z as the area of a rectangle and mark $d_x z$, $\Delta_y z$, Δz on the figure. Consider likewise $u = xyz$ as the volume of a rectangular parallelepiped.

16. *Small errors.* If $f(x, y)$ be a quantity determined by measurements on x and y, the error in f due to small errors dx, dy in x and y may be estimated as $df = f'_x dx + f'_y dy$ and the relative error may be taken as $df \div f = d\log f$. Why is this?

(α) Suppose $S = \frac{1}{2}ab\sin C$ be the area of a triangle with $a = 10$, $b = 20$, $C = 30°$. Find the error and the relative error if a is subject to an error of 0.1. *Ans.* 0.5, 1%.

(β) In (α) suppose C were liable to an error of 10′ of arc. *Ans.* 0.27, $\frac{1}{2}$%.

(γ) If a, b, C are liable to errors of 1%, the combined error in S may be 3.1%.

(δ) The radius r of a capillary tube is determined from $13.6\,\pi r^2 l = w$ by finding the weight w of a column of mercury of length l. If $w = 1$ gram with an error of 10^{-3} gr. and $l = 10$ cm. with an error of 0.2 cm., determine the possible error and relative error in r. *Ans.* 1.2%, 6×10^{-4}, mostly due to error in l.

(ϵ) The formula $c^2 = a^2 + b^2 - 2ab\cos C$ is used to determine c where $a = 20$, $b = 20$, $C = 60°$ with possible errors of 0.1 in a and b and 30′ in C. Find the possible absolute and relative errors in c. *Ans.* $\frac{1}{4}$, $1\frac{1}{4}$%.

(ζ) The possible percentage error of a product is the sum of the percentage errors of the factors.

(η) The constant g of gravity is determined from $g = 2st^{-2}$ by observing a body fall. If s is set at 4 ft. and t determined at about $\frac{1}{2}$ sec., show that the error in g is almost wholly due to the error in t, that is, that s can be set very much more accurately than t can be determined. For example, find the error in t which would make the same error in g as an error of $\frac{1}{8}$ inch in s.

(θ) The constant g is determined by $gt^2 = \pi^2 l$ with a pendulum of length l and period t. Suppose t is determined by taking the time 100 sec. of 100 beats of the pendulum with a stop watch that measures to $\frac{1}{5}$ sec. and that l may be measured as 100 cm. accurate to $\frac{1}{2}$ millimeter. Discuss the errors in g.

17. Let the coördinate x of a particle be $x = f(q_1, q_2)$ and depend on two independent variables q_1, q_2. Show that the velocity and kinetic energy are

$$v = f'_{q_1}\frac{dq_1}{dt} + f'_{q_2}\frac{dq_2}{dt}, \qquad T = \tfrac{1}{2}mv^2 = a_{11}\dot{q}_1^2 + 2a_{12}\dot{q}_1\dot{q}_2 + a_{22}\dot{q}_2^2,$$

where dots denote differentiation by t, and a_{11}, a_{12}, a_{22} are functions of (q_1, q_2). Show $\dfrac{\partial v}{\partial \dot{q}_i} = \dfrac{\partial x}{\partial q_i}$, $i = 1, 2$, and similarly for any number of variables q.

18. The helix $x = a\cos t$, $y = a\sin t$, $z = t\tan\alpha$ cuts the sphere $x^2 + y^2 + z^2 = a^2\sec^2\beta$ at $\sin^{-1}(\sin\alpha\sin\beta)$.

19. Apply the Theorem of the Mean to prove that $f(x, y, z)$ is a constant if $f'_x = f'_y = f'_z = 0$ is true for all values of x, y, z. Compare Theorem 16 (§ 27) and make the statement accurate.

20. Transform $\dfrac{df}{dn} = \sqrt{\left(\dfrac{\partial f}{\partial x}\right)^2 + \left(\dfrac{\partial f}{\partial y}\right)^2 + \left(\dfrac{\partial f}{\partial z}\right)^2}$ to (α) cylindrical and (β) polar coördinates (§ 40).

21. Find the angle of intersection of the helix $x = 2\cos t$, $y = 2\sin t$, $z = t$ and the surface $xyz = 1$ at their first intersection, that is, with $0 < t < \frac{1}{2}\pi$.

22. Let f, g, h be three functions of (x, y, z). In cylindrical coördinates (§ 40) form the combinations $F = f\cos\phi + g\sin\phi$, $G = -f\sin\phi + g\cos\phi$, $H = h$. Transform

$$(\alpha)\ \frac{\partial f}{\partial x} + \frac{\partial g}{\partial y} + \frac{\partial h}{\partial z}, \qquad (\beta)\ \frac{\partial h}{\partial y} - \frac{\partial g}{\partial z}, \qquad (\gamma)\ \frac{\partial g}{\partial x} - \frac{\partial f}{\partial y}$$

to cylindrical coördinates and express in terms of F, G, H in simplest form.

23. Given the functions y^x and $(z^y)^x$ and $z^{(y^x)}$. Find the total differentials and hence obtain the derivatives of x^x and $(x^x)^x$ and $x^{(x^x)}$.

50. Derivatives of higher order. If the first derivatives be again differentiated, there arise four derivatives f''_{xx}, f''_{xy}, f''_{yx}, f''_{yy} of the second order, where the first subscript denotes the first differentiation. These may also be written

$$f''_{xx} = \frac{\partial^2 f}{\partial x^2}, \quad f''_{xy} = \frac{\partial^2 f}{\partial y \partial x}, \quad f''_{yx} = \frac{\partial^2 f}{\partial x \partial y}, \quad f''_{yy} = \frac{\partial^2 f}{\partial y^2},$$

where the derivative of $\partial f/\partial y$ with respect to x is written $\partial^2 f/\partial x \partial y$ with the variables in the same order as required in $D_x D_y f$ and opposite to the order of the subscripts in f''_{yx}. This matter of order is usually of no importance owing to the theorem: *If the derivatives f'_x, f'_y have derivatives f''_{xy}, f''_{yx} which are continuous in (x, y) in the neighborhood of any point (x_0, y_0), the derivatives f''_{xy} and f''_{yx} are equal,* that is, $f''_{xy}(x_0, y_0) = f''_{yx}(x_0, y_0)$.

The theorem may be proved by repeated application of the Theorem of the Mean. For

$$[f(x_0 + h, y_0 + k) - f(x_0, y_0 + k)] - [f(x_0 + h, y_0) - f(x_0, y_0)] = [\phi(y_0 + k) - \phi(y_0)]$$
$$= [f(x_0 + h, y_0 + k) - f(x_0 + h, y_0)] - [f(x_0, y_0 + k) - f(x_0, y_0)] = [\psi(x_0 + h) - \psi(x_0)]$$

where $\phi(y)$ stands for $f(x_0 + h, y) - f(x_0, y)$ and $\psi(x)$ for $f(x, y_0 + k) - f(x, y_0)$. Now

$$\phi(y_0 + k) - \phi(y_0) = k\phi'(y_0 + \theta k) = k[f'_y(x_0 + h, y_0 + \theta k) - f'_y(x_0, y_0 + \theta k)],$$
$$\psi(x_0 + h) - \psi(x_0) = h\psi'(x_0 + \theta' h) = h[f'_x(x_0 + \theta' h, y_0 + k) - f'_x(x_0 + \theta' h, y_0)]$$

by applying the Theorem of the Mean to $\phi(y)$ and $\psi(x)$ regarded as functions of a single variable and then substituting. The results obtained are necessarily equal to each other; but each of these is in form for another application of the theorem.

$$k[f'_y(x_0+h,\ y_0+\theta k)-f'_y(x_0,\ y_0+\theta k)]=khf''_{yx}(x_0+\eta h,\ y_0+\theta k),$$
$$h[f'_x(x_0+\theta' h,\ y_0+k)-f'_x(x_0+\theta' h,\ y_0)]=hkf''_{xy}(x_0+\theta' h,\ y_0+\eta' k).$$

Hence $$f''_{yx}(x_0+\eta h,\ y_0+\theta k)=f''_{xy}(x_0+\theta' h,\ y_0+\eta' k).$$

As the derivatives f''_{yx}, f''_{xy} are supposed to exist and be continuous in the variables (x, y) at and in the neighborhood of (x_0, y_0), the limit of each side of the equation exists as $h \doteq 0$, $k \doteq 0$ and the equation is true in the limit. Hence

$$f''_{yx}(x_0, y_0)=f''_{xy}(x_0, y_0).$$

The differentiation of the three derivatives $f''_{xx}, f''_{xy}=f''_{yx}, f''_{yy}$ will give six derivatives of the third order. Consider f'''_{xxy} and f'''_{xyx}. These may be written as $(f'_x)''_{xy}$ and $(f'_x)''_{yx}$ and are equal by the theorem just proved (provided the restrictions as to continuity and existence are satisfied). A similar conclusion holds for f'''_{yxy} and f'''_{yyx}; the number of distinct derivatives of the third order reduces from six to four, just as the number of the second order reduces from four to three. In like manner for derivatives of any order, *the value of the derivative depends not on the order in which the individual differentiations with respect to x and y are performed, but only on the total number of differentiations with respect to each,* and the result may be written with the differentiations collected as

$$D_x^m D_y^n f=\frac{\partial^{m+n} f}{\partial x^m \partial y^n}=f^{(m+n)}_{x^m y^n}, \text{ etc.} \tag{22}$$

Analogous results hold for functions of any number of variables. If several derivatives are to be found and added together, a symbolic form of writing is frequently advantageous. For example,

$$(D_x^2 D_y D_z^3 + D_y^6)f=\frac{\partial^6 f}{\partial x^2 \partial y \partial z^3}+\frac{\partial^6 f}{\partial y^6}$$

or $$(D_x+D_y)^2 f=(D_x^2+2\,D_x D_y+D_y^2)f=f''_{xx}+2f''_{xy}+f''_{yy}.$$

51. It is sometimes necessary to *change the variable* in higher derivatives, particularly in those of the second order. This is done by a repeated application of (18). Thus f''_{rr} would be found by differentiating the first equation with respect to r, and f''_{rs} by differentiating the first by s or the second by r, and so on. Compare p. 12. The exercise below illustrates the method. It may be remarked that the use of *higher differentials* is often of advantage, although these differentials, like the higher differentials of functions of a single variable (Exs. 10, 16–19, p. 67), have the disadvantage that their form depends on what the independent variables are. This is also illustrated below. It should be particularly borne in mind that the great value of the first differential lies in the facts that it may be treated like a finite quantity and that its form is independent of the variables.

To change the variable in $v''_{xx} + v''_{yy}$ to polar coördinates and show

$$\frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} = \frac{\partial^2 v}{\partial r^2} + \frac{1}{r}\frac{\partial v}{\partial r} + \frac{1}{r^2}\frac{\partial^2 v}{\partial \phi^2}, \qquad \begin{cases} x = r\cos\phi, & y = r\sin\phi, \\ r = \sqrt{x^2+y^2}, & \phi = \tan^{-1}(y/x). \end{cases}$$

Then
$$\frac{\partial v}{\partial x} = \frac{\partial v}{\partial r}\frac{\partial r}{\partial x} + \frac{\partial v}{\partial \phi}\frac{\partial \phi}{\partial x}, \qquad \frac{\partial v}{\partial y} = \frac{\partial v}{\partial r}\frac{\partial r}{\partial y} + \frac{\partial v}{\partial \phi}\frac{\partial \phi}{\partial y}$$

by applying (18) directly with x, y taking the place of $r, s, \cdots$ and r, ϕ the place of $x, y, z, \cdots$. These expressions may be reduced so that

$$\frac{\partial v}{\partial x} = \frac{\partial v}{\partial r}\frac{x}{\sqrt{x^2+y^2}} + \frac{\partial v}{\partial \phi}\frac{-y}{x^2+y^2} = \frac{\partial v}{\partial r}\frac{x}{r} + \frac{\partial v}{\partial \phi}\frac{-y}{r^2}.$$

Next
$$\frac{\partial^2 v}{\partial x^2} = \frac{\partial}{\partial x}\frac{\partial v}{\partial x} = \frac{\partial}{\partial r}\frac{\partial v}{\partial x}\cdot\frac{\partial r}{\partial x} + \frac{\partial}{\partial \phi}\frac{\partial v}{\partial x}\cdot\frac{\partial \phi}{\partial x}$$
$$= \left[\frac{\partial^2 v}{\partial r^2}\frac{x}{r} + \frac{\partial v}{\partial r}\frac{\partial}{\partial r}\frac{x}{r} + \frac{\partial^2 v}{\partial r\partial\phi}\frac{-y}{r^2} + \frac{\partial v}{\partial \phi}\frac{\partial}{\partial r}\frac{-y}{r^2}\right]\frac{x}{r}$$
$$+ \left[\frac{\partial^2 v}{\partial\phi\partial r}\frac{x}{r} + \frac{\partial v}{\partial r}\frac{\partial}{\partial \phi}\frac{x}{r} + \frac{\partial^2 v}{\partial \phi^2}\frac{-y}{r^2} + \frac{\partial v}{\partial \phi}\frac{\partial}{\partial \phi}\frac{-y}{r^2}\right]\frac{-y}{r^2}.$$

The differentiations of x/r and $-y/r^2$ may be performed as indicated with respect to r, ϕ, remembering that, as r, ϕ are independent, the derivative of r by ϕ is 0. Then

$$\frac{\partial^2 v}{\partial x^2} = \frac{x^2}{r^2}\frac{\partial^2 v}{\partial r^2} + \frac{y^2}{r^3}\frac{\partial v}{\partial r} - 2\frac{xy}{r^3}\frac{\partial^2 v}{\partial r\partial\phi} + 2\frac{xy}{r^4}\frac{\partial v}{\partial \phi} + \frac{y^2}{r^4}\frac{\partial^2 v}{\partial \phi^2}.$$

In like manner $\partial^2 v/\partial y^2$ may be found, and the sum of the two derivatives reduces to the desired expression. This method is long and tedious though straightforward.

It is considerably shorter to start with the expression in polar coördinates and transform by the same method to the one in rectangular coördinates. Thus

$$\frac{\partial v}{\partial r} = \frac{\partial v}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial v}{\partial y}\frac{\partial y}{\partial r} = \frac{\partial v}{\partial x}\cos\phi + \frac{\partial v}{\partial y}\sin\phi = \frac{1}{r}\left(\frac{\partial v}{\partial x}x + \frac{\partial v}{\partial y}y\right),$$
$$\frac{\partial}{\partial r}\left(r\frac{\partial v}{\partial r}\right) = \left(\frac{\partial^2 v}{\partial x^2}\cos\phi + \frac{\partial^2 v}{\partial y\partial x}\sin\phi\right)x + \left(\frac{\partial^2 v}{\partial x\partial y}\cos\phi + \frac{\partial^2 v}{\partial y^2}\sin\phi\right)y + \frac{\partial v}{\partial x}\cos\phi + \frac{\partial v}{\partial y}\sin\phi,$$
$$\frac{\partial v}{\partial \phi} = \frac{\partial v}{\partial x}\frac{\partial x}{\partial \phi} + \frac{\partial v}{\partial y}\frac{\partial y}{\partial \phi} = -\frac{\partial v}{\partial x}r\sin\phi + \frac{\partial v}{\partial y}r\cos\phi = -\frac{\partial v}{\partial x}y + \frac{\partial v}{\partial y}x,$$
$$\frac{1}{r}\frac{\partial^2 v}{\partial \phi^2} = \left(\frac{\partial^2 v}{\partial x^2}\sin\phi - \frac{\partial^2 v}{\partial y\partial x}\cos\phi\right)y + \left(-\frac{\partial^2 v}{\partial x\partial y}\sin\phi + \frac{\partial^2 v}{\partial y^2}\cos\phi\right)x$$
$$-\frac{\partial v}{\partial x}\cos\phi - \frac{\partial v}{\partial y}\sin\phi.$$

Then
$$\frac{\partial}{\partial r}\left(r\frac{\partial v}{\partial r}\right) + \frac{1}{r}\frac{\partial^2 v}{\partial \phi^2} = \left(\frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2}\right)r$$

or
$$\frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} = \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial v}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2 v}{\partial \phi^2} = \frac{\partial^2 v}{\partial r^2} + \frac{1}{r}\frac{\partial v}{\partial r} + \frac{1}{r^2}\frac{\partial^2 v}{\partial \phi^2}. \tag{23}$$

The definitions $d_x^2 f = f''_{xx}dx^2$, $d_x d_y f = f''_{xy}dxdy$, $d_y^2 f = f''_{yy}dy^2$ would naturally be given for *partial differentials of the second order*, each of which would vanish if f reduced to either of the independent variables x, y or to any linear function of them. Thus the second differentials of the independent variables are zero. The second total differential would be obtained by differentiating the first total differential.

$$d^2f = ddf = d\left(\frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy\right) = d\frac{\partial f}{\partial x}dx + d\frac{\partial f}{\partial y}dy + \frac{\partial f}{\partial x}d^2x + \frac{\partial f}{\partial y}d^2y;$$

but $$d\frac{\partial f}{\partial x} = \frac{\partial^2 f}{\partial x^2}dx + \frac{\partial^2 f}{\partial y \partial x}dy, \qquad d\frac{\partial f}{\partial y} = \frac{\partial^2 f}{\partial x \partial y}dx + \frac{\partial^2 f}{\partial y^2}dy,$$

and $$d^2f = \frac{\partial^2 f}{\partial x^2}dx^2 + 2\frac{\partial^2 f}{\partial x \partial y}dxdy + \frac{\partial^2 f}{\partial y^2}dy^2 + \frac{\partial f}{\partial x}d^2x + \frac{\partial f}{\partial y}d^2y. \tag{24}$$

The last two terms vanish and the total differential reduces to the first three terms if x and y are the independent variables; and in this case the second derivatives, $f''_{xx}, f''_{xy}, f''_{yy}$, are the coefficients of dx^2, $2\,dxdy$, dy^2, which enables those derivatives to be found by an extension of the method of finding the first derivatives (§ 49). The method is particularly useful when all the second derivatives are needed.

The problem of the change of variable may now be treated. Let

$$\begin{aligned} d^2v &= \frac{\partial^2 v}{\partial x^2}dx^2 + 2\frac{\partial^2 v}{\partial x^2}dxdy + \frac{\partial^2 v}{\partial y^2}dy^2 \\ &= \frac{\partial^2 v}{\partial r^2}dr^2 + 2\frac{\partial^2 v}{\partial r \partial \phi}drd\phi + \frac{\partial^2 v}{\partial \phi^2}d\phi^2 + \frac{\partial v}{\partial r}d^2r + \frac{\partial v}{\partial \phi}d^2\phi, \end{aligned}$$

where x, y are the independent variables and r, ϕ other variables dependent on them — in this case, defined by the relations for polar coördinates. Then

$$dx = \cos\phi dr - r\sin\phi d\phi, \qquad dy = \sin\phi dr + r\cos\phi d\phi$$

or $$dr = \cos\phi dx + \sin\phi dy, \qquad rd\phi = -\sin\phi dx + \cos\phi dy. \tag{25}$$

Then $$d^2r = (-\sin\phi dx + \cos\phi dy)\,d\phi = rd\phi d\phi = rd\phi^2,$$

$$drd\phi + rd^2\phi = -(\cos\phi dx + \sin\phi dy)\,d\phi = -drd\phi,$$

where the differentials of dr and $rd\phi$ have been found subject to $d^2x = d^2y = 0$. Hence $d^2r = rd\phi^2$ and $rd^2\phi = -2\,drd\phi$. These may be substituted in d^2v which becomes

$$d^2v = \frac{\partial^2 v}{\partial r^2}dr^2 + 2\left(\frac{\partial^2 v}{\partial r \partial \phi} - \frac{1}{r}\frac{\partial v}{\partial \phi}\right)drd\phi + \left(\frac{\partial^2 v}{\partial \phi^2} + r\frac{\partial v}{\partial r}\right)d\phi^2.$$

Next the values of dr^2, $drd\phi$, $d\phi^2$ may be substituted from (25) and

$$\begin{aligned} d^2v &= \left[\frac{\partial^2 v}{\partial r^2}\cos^2\phi - \frac{2}{r}\left(\frac{\partial^2 v}{\partial r \partial \phi} - \frac{1}{r}\frac{\partial v}{\partial \phi}\right)\cos\phi\sin\phi + \left(\frac{\partial^2 v}{\partial \phi^2} + r\frac{\partial v}{\partial r}\right)\frac{\sin^2\phi}{r^2}\right]dx^2 \\ &+ 2\left[\frac{\partial^2 v}{\partial r^2}\cos\phi\sin\phi + \left(\frac{\partial^2 v}{\partial r \partial \phi} - \frac{1}{r}\frac{\partial v}{\partial \phi}\right)\frac{\cos^2\phi - \sin^2\phi}{r} - \frac{\partial^2 v}{\partial \phi^2}\frac{\cos\phi\sin\phi}{r^2}\right]dxdy \\ &+ \left[\frac{\partial^2 v}{\partial r^2}\sin^2\phi + \frac{2}{r}\left(\frac{\partial^2 v}{\partial r \partial \phi} - \frac{1}{r}\frac{\partial v}{\partial \phi}\right)\cos\phi\sin\phi + \left(\frac{\partial^2 v}{\partial \phi^2} + r\frac{\partial v}{\partial r}\right)\frac{\cos^2\phi}{r^2}\right]dy^2. \end{aligned}$$

Thus finally the derivatives v''_{xx}, v''_{xy}, v''_{yy} are the three brackets which are the coefficients of dx^2, $2\,dxdy$, dy^2. The value of $v''_{xx} + v''_{yy}$ is as found before.

52. The condition $f''_{xy} = f''_{yx}$ which subsists in accordance with the fundamental theorem of § 50 gives *the condition that*

$$M(x, y)\,dx + N(x, y)\,dy = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy = df$$

be the total differential of some function $f(x, y)$. *In fact*

$$\frac{\partial}{\partial y}\frac{\partial f}{\partial x} = \frac{\partial M}{\partial y} = \frac{\partial N}{\partial x} = \frac{\partial}{\partial x}\frac{\partial f}{\partial y}$$

and
$$\frac{\partial M}{\partial y} = \frac{\partial N}{\partial x} \quad \textit{or} \quad \left(\frac{dM}{dy}\right)_x = \left(\frac{dN}{dx}\right)_y. \tag{26}$$

The second form, where the variables which are constant during the differentiation are explicitly indicated as subscripts, is more common in works on thermodynamics. It will be proved later that conversely if this relation (26) holds, the expression $Mdx + Ndy$ is the total differential of some function, and the method of finding the function will also be given (§§ 92, 124). In case $Mdx + Ndy$ is the differential of some function $f(x, y)$ it is usually called an *exact differential.*

The application of the condition for an exact differential may be made in connection with a problem in thermodynamics. Let S and U be the entropy and energy of a gas or vapor inclosed in a receptacle of volume v and subjected to the pressure p at the temperature T. The fundamental equation of thermodynamics, connecting the differentials of energy, entropy, and volume, is

$$dU = TdS - pdv; \quad \text{and} \quad \left(\frac{dT}{dv}\right)_S = -\left(\frac{dp}{dS}\right)_v \tag{27}$$

is the condition that dU be a total differential. Now, any two of the five quantities U, S, v, T, p may be taken as independent variables. In (27) the choice is S, v; if the equation were solved for dS, the choice would be U, v; and U, S if solved for dv. In each case the cross differentiation to express the condition (26) would give rise to a relation between the derivatives.

If p, T were desired as independent variables, the change of variable

$$dS = \left(\frac{dS}{dp}\right)_T dp + \left(\frac{dS}{dT}\right)_p dT, \qquad dv = \left(\frac{dv}{dp}\right)_T dp + \left(\frac{dv}{dT}\right)_p dT$$

with
$$dU = \left[T\left(\frac{dS}{dp}\right)_T - p\left(\frac{dv}{dp}\right)_T\right]dp + \left[T\left(\frac{dS}{dT}\right)_p - p\left(\frac{dv}{dT}\right)_p\right]dT$$

should be made. The expression of the condition is then

$$\left\{\frac{d}{dT}\left[T\left(\frac{dS}{dp}\right)_T - p\left(\frac{dv}{dp}\right)_T\right]\right\}_p = \left\{\frac{d}{dp}\left[T\left(\frac{dS}{dT}\right)_p - p\left(\frac{dv}{dT}\right)_p\right]\right\}_T$$

or
$$\left(\frac{dS}{dp}\right)_T + T\frac{\partial S}{\partial T\partial p} - p\frac{\partial^2 v}{\partial T\partial p} = T\frac{\partial^2 S}{\partial p\partial T} - \left(\frac{dv}{dT}\right)_p - p\frac{\partial^2 v}{\partial p\partial T},$$

where the differentiation on the left is made with p constant and that on the right with T constant and where the subscripts have been dropped from the second derivatives and the usual notation adopted. Everything cancels except two terms which give

$$\left(\frac{dS}{dp}\right)_T = -\left(\frac{dv}{dT}\right)_p \quad \text{or} \quad \frac{1}{T}\left(\frac{TdS}{dp}\right)_T = -\left(\frac{dv}{dT}\right)_p. \tag{28}$$

The importance of the test for an exact differential lies not only in the relations obtained between the derivatives as above, but also in the fact that in applied mathematics a great many expressions are written as differentials which are not the total differentials of any functions and which must be distinguished from exact differentials. For instance if dH denote the infinitesimal portion of heat added to the gas or vapor above considered, the fundamental equation is expressed as $dH = dU + pdv$. That is to say, the amount of heat added is equal to the increase in the energy plus the work done by the gas in expanding. Now dH is not the differential of any function $H(U, v)$; it is $dS = dH/T$ which is the differential, and this is one reason for introducing the entropy S. Again if the forces X, Y act on a particle, the *work* done during the displacement through the arc $ds = \sqrt{dx^2 + dy^2}$ is written $dW = Xdx + Ydy$. It may happen that this is the total differential of some function; indeed, if

$$dW = -dV(x, y), \quad Xdx + Ydy = -dV, \quad X = -\frac{\partial V}{\partial x}, \quad Y = -\frac{\partial V}{\partial y},$$

where the negative sign is introduced in accordance with custom, the function V is called the *potential energy* of the particle. In general, however, there is no potential energy function V, and dW is not an exact differential; this is always true when part of the work is due to forces of friction. A notation which should distinguish between exact differentials and those which are not exact is much more needed than a notation to distinguish between partial and ordinary derivatives; but there appears to be none.

Many of the physical magnitudes of thermodynamics are expressed as derivatives and such relations as (26) establish relations between the magnitudes. Some definitions:

specific heat at constant volume is $C_v = \left(\frac{dH}{dT}\right)_v = T\left(\frac{dS}{dT}\right)_v$,

specific heat at constant pressure is $C_p = \left(\frac{dH}{dT}\right)_p = T\left(\frac{dS}{dT}\right)_p$,

latent heat of expansion is $L_v = \left(\frac{dH}{dv}\right)_T = T\left(\frac{dS}{dv}\right)_T$,

coefficient of cubic expansion is $\alpha_p = \frac{1}{v}\left(\frac{dv}{dT}\right)_p$,

modulus of elasticity (isothermal) is $E_T = -v\left(\frac{dp}{dv}\right)_T$,

modulus of elasticity (adiabatic) is $E_S \doteq -v\left(\frac{dp}{dv}\right)_S$.

53. A polynomial is said to be homogeneous when each of its terms is of the same order when all the variables are considered. A definition of homogeneity which includes this case and is applicable to more general cases is: *A function $f(x, y, z, \cdots)$ of any number of variables is called homogeneous if the function is multiplied by some power of λ when all the variables are multiplied by λ*; and the power of λ which factors out is called the order of homogeneity of the function. In symbols the condition for homogeneity of order n is

$$f(\lambda x, \lambda y, \lambda z, \cdots) = \lambda^n f(x, y, z, \cdots). \tag{29}$$

Thus

$$xe^{\frac{y}{x}} + \frac{y^2}{x}, \quad \frac{xy}{z^2} + \tan^{-1}\frac{x}{z}, \quad \frac{1}{\sqrt{x^2+y^2}} \tag{29'}$$

are homogeneous functions of order 1, 0, -1 respectively. To test a function for homogeneity it is merely necessary to replace all the variables by λ times the variables and see if λ factors out completely. The homogeneity may usually be seen without the test.

If the identity (29) be differentiated with respect to λ,

$$\left(x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y} + z\frac{\partial}{\partial z} + \cdots\right)f(\lambda x, \lambda y, \lambda z, \cdots) = n\lambda^{n-1}f(x, y, z, \cdots).$$

A second differentiation with respect to λ would give

$$\left(x^2\frac{\partial^2}{\partial x^2} + xy\frac{\partial^2}{\partial x\partial y} + xz\frac{\partial^2}{\partial x\partial z} + \cdots\right)f + \left(yx\frac{\partial^2}{\partial y\partial x} + y^2\frac{\partial^2}{\partial y^2} + yz\frac{\partial^2}{\partial z^2} + \cdots\right)f$$

$$+\left(zx\frac{\partial^2}{\partial z\partial x} + zy\frac{\partial^2}{\partial z\partial y} + z^2\frac{\partial^2}{\partial z^2} + \cdots\right)f + \cdots = n(n-1)\lambda^{n-2}f(x, y, z, \cdots)$$

or

$$\left(x^2\frac{\partial^2}{\partial x^2} + 2xy\frac{\partial^2}{\partial x\partial y} + y^2\frac{\partial^2}{\partial y^2} + \cdots\right)f = n(n-1)\lambda^{n-2}f(x, y, z, \cdots).$$

Now if λ be set equal to 1 in these equations, then

$$x\frac{\partial f}{\partial x} + y\frac{\partial f}{\partial y} + z\frac{\partial f}{\partial z} + \cdots = nf(x, y, z, \cdots), \tag{30}$$

$$x^2\frac{\partial^2 f}{\partial x^2} + 2xy\frac{\partial^2 f}{\partial x\partial y} + y^2\frac{\partial^2 f}{\partial y^2} + 2xz\frac{\partial^2 f}{\partial x\partial z} + \cdots = n(n-1)f(x, y, z, \cdots).$$

In words, these equations state that the sum of the partial derivatives each multiplied by the variable with respect to which the differentiation is performed is n times the function if the function is homogeneous of order n; and that the sum of the second derivatives each multiplied by the variables involved and by 1 or 2, according as the variable is repeated or not, is $n(n-1)$ times the function. The general formula obtained by differentiating any number of times with respect to λ may be expressed symbolically in the convenient form

$$(xD_x + yD_y + zD_z + \cdots)^k f = n(n-1)\cdots(n-k+1)f. \tag{31}$$

This is known as *Euler's Formula* on homogeneous functions.

It is worth while noting that in a certain sense every equation which represents a geometric or physical relation is homogeneous. For instance, in geometry the magnitudes that arise may be lengths, areas, volumes, or angles. These magnitudes are expressed as a number times a unit; thus, $\sqrt{2}$ ft., 3 sq. yd., π cu. ft.

In adding and subtracting, the terms must be like quantities; lengths added to lengths, areas to areas, etc. The *fundamental unit* is taken as length. The units of area, volume, and angle are *derived* therefrom. Thus the area of a rectangle or the volume of a rectangular parallelepiped is

$$A = a \text{ ft.} \times b \text{ ft.} = ab \text{ ft.}^2 = ab \text{ sq ft.}, \quad V = a \text{ ft.} \times b \text{ ft.} \times c \text{ ft.} = abc \text{ ft.}^3 = abc \text{ cu. ft.},$$

and the units sq. ft., cu. ft. are denoted as ft.^{2}, ft.^{3} just as if the simple unit ft. had been treated as a literal quantity and included in the multiplication. An area or volume is therefore considered as a compound quantity consisting of a number which gives its magnitude and a unit which gives its quality or dimensions. If L denote length and $[L]$ denote "of the dimensions of length," and if similar notations be introduced for area and volume, the equations $[A] = [L]^2$ and $[V] = [L]^3$ state that the dimensions of area are squares of length, and of volumes, cubes of lengths. If it be recalled that for purposes of analysis an angle is measured by the ratio of the arc subtended to the radius of the circle, the dimensions of angle are seen to be nil, as the definition involves the ratio of like magnitudes and must therefore be a *pure number*.

When geometric facts are represented analytically, either of two alternatives is open: 1°, the equations may be regarded as existing between mere numbers; or 2°, as between actual magnitudes. Sometimes one method is preferable, sometimes the other. Thus the equation $x^2 + y^2 = r^2$ of a circle may be interpreted as 1°, the sum of the squares of the coördinates (numbers) is constant; or 2°, the sum of the squares on the legs of a right triangle is equal to the square on the hypotenuse (Pythagorean Theorem). The second interpretation better sets forth the true inwardness of the equation. Consider in like manner the parabola $y^2 = 4px$. Generally y and x are regarded as mere numbers, but they may equally be looked upon as lengths and then the statement is that the square upon the ordinate equals the rectangle upon the abscissa and the constant length $4p$; this may be interpreted into an actual construction for the parabola, because a square equivalent to a rectangle may be constructed.

In the last interpretation the constant p was assigned the dimensions of length so as to render the equation homogeneous in dimensions, with each term of the dimensions of area or $[L]^2$. It will be recalled, however, that in the definition of the parabola, the quantity p actually has the dimensions of length, being half the distance from the fixed point to the fixed line (focus and directrix). This is merely another corroboration of the initial statement that the equations which actually arise in considering geometric problems are homogeneous in their dimensions, and must be so for the reason that in stating the first equation like magnitudes must be compared with like magnitudes.

The question of dimensions may be carried along through such processes as differentiation and integration. For let y have the dimensions $[y]$ and x the dimensions $[x]$. Then Δy, the difference of two y's, must still have the dimensions $[y]$ and Δx the dimensions $[x]$. The quotient $\Delta y/\Delta x$ then has the dimensions $[y]/[x]$. For example the relations for area and for volume of revolution,

$$\frac{dA}{dx} = y, \quad \frac{dV}{dx} = \pi y^2, \quad \text{give} \quad \left[\frac{dA}{dx}\right] = \frac{[A]}{[L]} = [L], \quad \left[\frac{dV}{dx}\right] = \frac{[V]}{[L]} = [L]^2,$$

and the dimensions of the left-hand side check with those of the right-hand side. As integration is the limit of a sum, the dimensions of an integral are the product of the dimensions of the function to be integrated and of the differential dx. Thus if

$$y = \int_0^x \frac{dx}{a^2 + x^2} = \frac{1}{a} \tan^{-1} \frac{x}{a} + c$$

were an integral arising in actual practice, the very fact that a^2 and x^2 are added would show that they must have the same dimensions. If the dimensions of x be $[L]$, then

$$\left[\int_0^x \frac{dx}{a^2 + x^2}\right] = \left[\frac{1}{a^2 + x^2}\right][dx] = \frac{1}{[L]^2}[L] = \frac{1}{[L]} = [y],$$

and this checks with the dimensions on the right which are $[L]^{-1}$, since angle has no dimensions. As a rule, the theory of dimensions is neglected in pure mathematics; but it can nevertheless be made exceedingly useful and instructive.

In mechanics the *fundamental units* are length, mass, and time; and are denoted by $[L]$, $[M]$, $[T]$. The following table contains some derived units:

velocity	$\frac{[L]}{[T]}$,	acceleration	$\frac{[L]}{[T]^2}$,	force	$\frac{[M][L]}{[T]^2}$,
areal velocity	$\frac{[L]^2}{[T]}$,	density	$\frac{[M]}{[L]^3}$,	momentum	$\frac{[M][L]}{[T]}$,
angular velocity	$\frac{1}{[T]}$,	moment	$\frac{[M][L]^2}{[T]^2}$,	energy	$\frac{[M][L]^2}{[T]^2}$.

With the aid of a table like this it is easy to convert magnitudes in one set of units as ft., lb., sec., to another system, say cm., gm., sec. All that is necessary is to substitute for each individual unit its value in the new system. Thus

$$g = 32\tfrac{1}{6}\,\frac{\text{ft.}}{\text{sec.}^2}, \qquad 1 \text{ ft.} = 30.48 \text{ cm.}, \qquad g = 32\tfrac{1}{6} \times 30.48\,\frac{\text{cm.}}{\text{sec.}^2} = 980\tfrac{1}{2}\,\frac{\text{cm.}}{\text{sec.}^2}.$$

EXERCISES

1. Obtain the derivatives f''_{xx}, f''_{xy}, f''_{yx}, f''_{yy} and verify $f''_{xy} = f''_{yx}$.

$$(\alpha)\ \sin^{-1}\frac{y}{x}, \qquad (\beta)\ \log\frac{x^2 + y^2}{xy}, \qquad (\gamma)\ \phi\left(\frac{y}{x}\right) + \psi(xy).$$

2. Compute $\partial^2 v/\partial y^2$ in polar coördinates by the straightforward method.

3. Show that $a^2 \dfrac{\partial^2 v}{\partial x^2} = \dfrac{\partial^2 v}{\partial t^2}$ if $v = f(x + at) + \phi(x - at)$.

4. Show that this equation is unchanged in form by the transformation:

$$\frac{\partial^2 f}{\partial x^2} + 2xy^2 \frac{\partial f}{\partial x} + 2(y - y^3)\frac{\partial f}{\partial y} + x^2y^2 f = 0; \qquad u = xy, \quad v = 1/y.$$

5. In polar coördinates $z = r\cos\theta$, $x = r\sin\theta\cos\phi$, $y = r\sin\theta\sin\phi$ in space

$$\frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} + \frac{\partial^2 v}{\partial y^2} = \frac{1}{r^2}\left[\frac{\partial}{\partial r}\left(r^2 \frac{\partial v}{\partial r}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2 v}{\partial \phi^2} + \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\,\frac{\partial v}{\partial\theta}\right)\right].$$

The work of transformation may be shortened by substituting successively

$$x = r_1\cos\phi, \quad y = r_1\sin\phi, \quad \text{and} \quad z = r\cos\phi, \quad r_1 = r\sin\phi.$$

6. Let x, y, z, t be four independent variables and $x = r\cos\phi$, $y = r\sin\phi$, $z = z$ the equations for transforming x, y, z to cylindrical coördinates. Let

$$X = -\frac{\partial^2 f}{\partial x \partial z}, \quad Y = -\frac{\partial^2 f}{\partial y \partial z}, \quad Z = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}, \quad F = \frac{\partial^2 f}{\partial y \partial t}, \quad G = -\frac{\partial^2 f}{\partial x \partial t};$$

show $Z = \frac{1}{r}\frac{\partial Q}{\partial r}, \quad X\cos\phi + Y\sin\phi = -\frac{1}{r}\frac{\partial Q}{\partial z}, \quad F\sin\phi - G\cos\phi = \frac{1}{r}\frac{\partial Q}{\partial t},$

where $r^{-1}Q = \partial f/\partial r$. (Of importance for the Hertz oscillator.)

7. Apply the test for an exact differential to each of the following, and write by inspection the functions corresponding to the exact differentials:

(α) $3xdx + y^2dy$, $\quad(\beta)$ $3xydx + x^3dy$, $\quad(\gamma)$ $x^2ydx + y^2dy$,

(δ) $\dfrac{xdx + ydy}{x^2 + y^2}$, $\quad(\epsilon)$ $\dfrac{xdx - ydy}{x^2 + y^2}$, $\quad(\zeta)$ $\dfrac{ydx - xdy}{x^2 + y^2}$,

(η) $(4x^3 + 3x^2y + y^2)dx + (x^3 + 2xy + 3y^3)dy$, $\quad(\theta)$ $x^2y^2(dx + dy)$.

8. Express the conditions that $P(x, y, z)dx + Q(x, y, z)dy + R(x, y, z)dz$ be an exact differential $dF(x, y, z)$. Apply these conditions to the differentials:

(α) $3x^2y^2zdx + 2x^3yzdy + x^3y^2dz$, $\quad(\beta)$ $(y + z)dx + (x + z)dy + (x + y)dz$.

9. Obtain $\left(\dfrac{dp}{dT}\right)_v = \left(\dfrac{dS}{dv}\right)_T$ and $\left(\dfrac{dv}{dS}\right)_p = \left(\dfrac{dT}{dp}\right)_S$ from (27) with proper variables.

10. If three functions (called thermodynamic potentials) be defined as

$$\psi = U - TS, \qquad \chi = U + pv, \qquad \zeta = U - TS + pv,$$

show $\quad d\psi = -SdT - pdv, \quad d\chi = TdS + vdp, \quad d\zeta = -SdT + vdp,$

and express the conditions that $d\psi$, $d\chi$, $d\zeta$ be exact. Compare with Ex. 9.

11. State in words the definitions corresponding to the defining formulas, p. 107.

12. If the sum $(Mdx + Ndy) + (Pdx + Qdy)$ of two differentials is exact and one of the differentials is exact, the other is. Prove this.

13. Apply Euler's Formula (31), for the simple case $k = 1$, to the three functions (29′) and verify the formula. Apply it for $k = 2$ to the first function.

14. Verify the homogeneity of these functions and determine their order:

(α) $y^2/x + x(\log x - \log y)$, $\quad(\beta)$ $\dfrac{x^m y^n}{\sqrt{x^2 + y^2}}$, $\quad(\gamma)$ $\dfrac{xyz}{ax + by + cz}$,

(δ) $xye^{\frac{xz}{y^2}} + z^2$, $\quad(\epsilon)$ $\sqrt{x}\cot^{-1}\dfrac{y}{z}$, $\quad(\zeta)$ $\dfrac{\sqrt[5]{x} - \sqrt[5]{y}}{\sqrt[4]{x} + \sqrt[4]{y}}$.

15. State the dimensions of moment of inertia and convert a unit of moment of inertia in ft.-lb. into its equivalent in cm.-gm.

16. Discuss for dimensions Peirce's formulas Nos. 93, 124–125, 220, 300.

17. Continue Ex. 17, p. 101, to show $\dfrac{d}{dt}\dfrac{\partial x}{\partial q_i} = \dfrac{\partial v}{\partial q_i}$ and $\dfrac{d}{dt}\dfrac{\partial T}{\partial \dot q_i} = m\dot v\dfrac{\partial x}{\partial q_i} + \dfrac{\partial T}{\partial q_i}$.

18. If $p_i = \dfrac{\partial T}{\partial \dot q_i}$ in Ex. 17, p. 101, show without analysis that $2T = \dot q_1 p_1 + \dot q_2 p_2$. If T' denote $T' = T$, where T' is considered as a function of p_1, p_2 while T is considered as a function of $\dot q_1$, $\dot q_2$, prove from $T' = \dot q_1 p_1 + \dot q_2 p_2 - T$ that

$$\frac{\partial T'}{\partial p_i} = \dot q_i, \qquad \frac{\partial T'}{\partial q_i} = -\frac{\partial T}{\partial q_i}.$$

19. If (x_1, y_1) and (x_2, y_2) are the coördinates of two moving particles and

$$m_1\frac{d^2x_1}{dt^2}=X_1,\quad m_1\frac{d^2y_1}{dt^2}=Y_1,\quad m_2\frac{d^2x_2}{dt^2}=X_2,\quad m_2\frac{d^2y_2}{dt^2}=Y_2$$

are the equations of motion, and if x_1, y_1, x_2, y_2 are expressible as

$$x_1=f_1(q_1, q_2, q_3),\quad y_1=g_1(q_1, q_2, q_3),\quad x_2=f_2(q_1, q_2, q_3),\quad y_2=g_2(q_1, q_2, q_3)$$

in terms of three independent variables q_1, q_2, q_3, show that

$$Q_1=X_1\frac{\partial x_1}{\partial q_1}+Y_1\frac{\partial y_1}{\partial q_1}+X_2\frac{\partial x_2}{\partial q_1}+Y_2\frac{\partial y_2}{\partial q_1}=\frac{d}{dt}\frac{\partial T}{\partial \dot q_1}-\frac{\partial T}{\partial q_1},$$

where $T=\frac{1}{2}(m_1v_1^2+m_2v_2^2)=T(q_1, q_2, q_3, \dot q_1, \dot q_2, \dot q_3)$ and is homogeneous of the second degree in $\dot q_1, \dot q_2, \dot q_3$. The work may be carried on as a generalization of Ex. 17, p. 101, and Ex. 17 above. It may be further extended to any number of particles whose positions in space depend on a number of variables q.

20. In Ex. 19 if $p_i=\dfrac{\partial T}{\partial \dot q_i}$, generalize Ex. 18 to obtain

$$\dot q_i=\frac{\partial T'}{\partial p_i},\quad \frac{\partial T'}{\partial q_i}=-\frac{\partial T}{\partial q_i},\quad Q_1=\frac{dp_1}{dt}+\frac{\partial T'}{\partial q_1}.$$

The equations $Q_i=\dfrac{d}{dt}\dfrac{\partial T}{\partial \dot q_i}-\dfrac{\partial T}{\partial q_i}$ and $Q_i=\dfrac{dp_i}{dt}+\dfrac{\partial T'}{\partial q_i}$ are respectively the Lagrangian and Hamiltonian equations of motion.

21. If $rr'=k^2$ and $\phi'=\phi$ and $v'(r', \phi')=v(r, \phi)$, show

$$\frac{\partial^2 v'}{\partial r'^2}+\frac{1}{r'}\frac{\partial v'}{\partial r'}+\frac{1}{r'^2}\frac{\partial^2 v'}{\partial \phi'^2}=\frac{r^2}{r'^2}\left(\frac{\partial^2 v}{\partial r^2}+\frac{1}{r}\frac{\partial v}{\partial r}+\frac{1}{r^2}\frac{\partial^2 v}{\partial \phi^2}\right).$$

22. If $rr'=k^2$, $\phi'=\phi$, $\theta'=\theta$, and $v'(r', \phi', \theta')=\dfrac{k}{r'}v(r, \phi, \theta)$, show that the expression of Ex. 5 in the primed letters is kr^2/r'^3 of its value for the unprimed letters. (Useful in § 198.)

23. If $z=x\phi\left(\dfrac{y}{x}\right)+\psi\left(\dfrac{y}{x}\right)$, show $x^2\dfrac{\partial^2 z}{\partial x^2}+2xy\dfrac{\partial^2 z}{\partial x\partial y}+y^2\dfrac{\partial^2 z}{\partial y^2}=0.$

24. Make the indicated changes of variable:

(α) $\dfrac{\partial^2 V}{\partial x^2}+\dfrac{\partial^2 V}{\partial y^2}=e^{-2u}\left(\dfrac{\partial^2 V}{\partial u^2}+\dfrac{\partial^2 V}{\partial v^2}\right)$ if $x=e^u\cos v$, $y=e^u\sin v$,

(β) $\dfrac{\partial^2 V}{\partial u^2}+\dfrac{\partial^2 V}{\partial v^2}=\left(\dfrac{\partial^2 V}{\partial x^2}+\dfrac{\partial^2 V}{\partial y^2}\right)\left[\left(\dfrac{\partial f}{\partial u}\right)^2+\left(\dfrac{\partial f}{\partial v}\right)^2\right]$, where

$$x=f(u, v),\quad y=\phi(u, v),\quad \frac{\partial f}{\partial u}=\frac{\partial \phi}{\partial v},\quad \frac{\partial f}{\partial v}=-\frac{\partial \phi}{\partial u}.$$

25. For an orthogonal transformation (Ex. 10 (ζ), p. 100)

$$\frac{\partial^2 v}{\partial x^2}+\frac{\partial^2 v}{\partial y^2}+\frac{\partial^2 v}{\partial z^2}=\frac{\partial^2 v}{\partial x'^2}+\frac{\partial^2 v}{\partial y'^2}+\frac{\partial^2 v}{\partial z'^2}.$$

54. Taylor's Formula and applications. The development of $f(x, y)$ is found, as was the Theorem of the Mean, from the relation (p. 95)

$$\Delta f = \Phi(1) - \Phi(0) \quad \text{if} \quad \Phi(t) = f(a + th,\ b + tk).$$

If $\Phi(t)$ be expanded by Maclaurin's Formula to n terms,

$$\Phi(t) - \Phi(0) = t\Phi'(0) + \frac{t^2}{2!}\Phi''(0) + \cdots + \frac{t^{n-1}}{(n-1)!}\Phi^{(n-1)}(0) + \frac{t^n}{n!}\Phi^{(n)}(\theta t).$$

The expressions for $\Phi'(t)$ and $\Phi'(0)$ may be found as follows by (10):

$$\Phi'(t) = hf'_x + kf'_y, \quad \Phi'(0) = [hf'_x + kf'_y]_{\substack{x=a\\y=b}},$$

then
$$\begin{aligned}\Phi''(t) &= h(hf''_{xx} + kf''_{yx}) + k(hf''_{xy} + kf''_{yy})\\ &= h^2f''_{xx} + 2hkf''_{xy} + k^2f''_{yy} = (hD_x + kD_y)^2 f,\end{aligned}$$

$$\Phi^{(i)}(t) = (hD_x + kD_y)^i f, \quad \Phi^{(i)}(0) = [(hD_x + kD_y)^i f]_{\substack{x=a\\y=b}}.$$

And
$$\begin{aligned}f(a+h, b+k) - f(a, b) &= \Delta f = \Phi(1) - \Phi(0) = (hD_x + kD_y)f(a, b)\\ &+ \frac{1}{2!}(hD_x + kD_y)^2 f(a, b) + \cdots + \frac{1}{(n-1)!}(hD_x + kD_y)^{n-1}f(a, b)\\ &+ \frac{1}{n!}(hD_x + kD_y)^n f(a + \theta h, b + \theta k).\end{aligned} \tag{32}$$

In this expansion, the increments h and k may be replaced, if desired, by $x - a$ and $y - b$ and then $f(x, y)$ will be expressed in terms of its value and the values of its derivatives at (a, b) in a manner entirely analogous to the case of a single variable. In particular if the point (a, b) about which the development takes place be $(0, 0)$ the development becomes Maclaurin's Formula for $f(x, y)$.

$$\begin{aligned}f(x, y) &= f(0, 0) + (xD_x + yD_y)f(0, 0) + \frac{1}{2!}(xD_x + yD_y)^2 f(0, 0) + \cdots\\ &+ \frac{1}{(n-1)!}(xD_x + yD_y)^{n-1}f(0, 0) + \frac{1}{n!}(xD_x + yD_y)^n f(\theta x, \theta y).\end{aligned} \tag{32'}$$

Whether in Maclaurin's or Taylor's Formula, the successive terms are homogeneous polynomials of the 1st, 2d, $\cdots$, $(n-1)$st order in x, y or in $x - a$, $y - b$. The formulas are unique as in § 32.

Suppose $\sqrt{1 - x^2 - y^2}$ is to be developed about $(0, 0)$. The successive derivatives are

$$f'_x = \frac{-x}{\sqrt{1 - x^2 - y^2}}, \quad f'_y = \frac{-y}{\sqrt{1 - x^2 - y^2}}, \quad f'_x(0, 0) = 0, \quad f'_y(0, 0) = 0,$$

$$f''_{xx} = \frac{-1 + y^2}{(1 - x^2 - y^2)^{\frac{3}{2}}}, \quad f''_{xy} = \frac{xy}{(1 - x^2 - y^2)^{\frac{3}{2}}}, \quad f''_{yy} = \frac{-1 + x^2}{(1 - x^2 - y^2)^{\frac{3}{2}}},$$

$$f'''_{x^3} = \frac{\frac{3}{2}(1 - y^2)x}{(1 - x^2 - y^2)^{\frac{5}{2}}}, \quad f'''_{x^2y} = \frac{y^3 - 2xy^2 - y}{(1 - x^2 - y^2)^{\frac{5}{2}}}, \quad \cdots,$$

and $\sqrt{1 - x^2 - y^2} = 1 + (0x + 0y) + \frac{1}{2}(-x^2 + 0xy - y^2) + \frac{1}{6}(0x^3 + \cdots) + \cdots,$

or $\sqrt{1 - x^2 - y^2} = 1 - \frac{1}{2}(x^2 + y^2) +$ terms of fourth order $+ \cdots$.

In this case the expansion may be found by treating $x^2 + y^2$ as a single term and expanding by the binomial theorem. The result would be

$$[1-(x^2+y^2)]^{\frac{1}{2}} = 1 - \tfrac{1}{2}(x^2+y^2) - \tfrac{1}{8}(x^4+2x^2y^2+y^4) - \tfrac{1}{16}(x^2+y^2)^3 - \cdots.$$

That the development thus obtained is identical with the Maclaurin development that might be had by the method above, follows from the uniqueness of the development. Some such short cut is usually available.

55. The condition that a function $z = f(x, y)$ have a minimum or maximum at (a, b) is that $\Delta f > 0$ or $\Delta f < 0$ for all values of $h = \Delta x$ and $k = \Delta y$ which are sufficiently small. From either geometrical or analytic considerations it is seen that if the surface $z = f(x, y)$ has a minimum or maximum at (a, b), the curves in which the planes $y = b$ and $x = a$ cut the surface have minima or maxima at $x = a$ and $y = b$ respectively. Hence the partial derivatives f'_x and f'_y must both vanish at (a, b), provided, of course, that exceptions like those mentioned on page 7 be made. The two simultaneous equations

$$f'_x = 0, \quad f'_y = 0, \tag{33}$$

corresponding to $f'(x) = 0$ in the case of a function of a single variable, may then be solved to find the positions (x, y) of the minima and maxima. Frequently the geometric or physical interpretation of $z = f(x, y)$ or some special device will then determine whether there is a maximum or a minimum or neither at each of these points.

For example let it be required to find the maximum rectangular parallelepiped which has three faces in the coördinate planes and one vertex in the plane $x/a + y/b + z/c = 1$. The volume is

$$V = xyz = cxy\left(1 - \frac{x}{a} - \frac{y}{b}\right).$$

$$\frac{\partial V}{\partial x} = -2\frac{c}{a}xy - \frac{c}{b}y^2 + cy = 0 \qquad \frac{\partial V}{\partial y} = -2\frac{c}{b}xy - \frac{c}{a}x^2 + cx = 0.$$

The solution of these equations is $x = \frac{1}{3}a$, $y = \frac{1}{3}b$. The corresponding z is $\frac{1}{3}c$ and the volume V is therefore $abc/9$ or $\frac{2}{9}$ of the volume cut off from the first octant by the plane. It is evident that this solution is a maximum. There are other solutions of $V'_x = V'_y = 0$ which have been discarded because they give $V = 0$.

The conditions $f'_x = f'_y = 0$ may be established analytically. For

$$\Delta f = (f'_x + \zeta_1)\Delta x + (f'_y + \zeta_2)\Delta y.$$

Now as ζ_1, ζ_2 are infinitesimals, the signs of the parentheses are determined by the signs of f'_x, f'_y unless these derivatives vanish; and hence unless $f'_x = 0$, the sign of Δf for Δx sufficiently small and positive and $\Delta y = 0$ would be opposite to the sign of Δf for Δx sufficiently small and negative and $\Delta y = 0$. Therefore *for a minimum or maximum* $f'_x = 0$; *and in like manner* $f'_y = 0$. Considerations like these will serve to establish a criterion for distinguishing between maxima and minima analogous to the criterion furnished by $f''(x)$ in the case of one variable. For if $f'_x = f'_y = 0$, then

$$\Delta f = \tfrac{1}{2}(h^2 f''_{xx} + 2\,hk f''_{xy} + k^2 f''_{yy})_{x=a+\theta h,\,y=b+\theta k},$$

by Taylor's Formula to two terms. Now if the second derivatives are continuous functions of (x, y) in the neighborhood of (a, b), each derivative at $(a+\theta h, b+\theta k)$ may be written as its value at (a, b) plus an infinitesimal. Hence

$$\Delta f = \tfrac{1}{2}(h^2 f''_{xx} + 2\,hk f''_{xy} + k^2 f''_{yy})_{(a,\,b)} + \tfrac{1}{2}(h^2\zeta_1 + 2\,hk\zeta_2 + k^2\zeta_3).$$

Now the sign of Δf for sufficiently small values of h, k must be the same as the sign of the first parenthesis provided that parenthesis does not vanish. Hence if the quantity

$$(h^2 f''_{xx} + 2\,hk f''_{xy} + k^2 f''_{yy})_{(a,\,b)} \begin{matrix} > 0 \text{ for every } (h, k), \text{ a minimum} \\ < 0 \text{ for every } (h, k), \text{ a maximum.} \end{matrix}$$

As the derivatives are taken at the point (a, b), they have certain constant values, say A, B, C. The question of distinguishing between minima and maxima therefore reduces to the discussion of the possible signs of a *quadratic form* $Ah^2 + 2\,Bhk + Ck^2$ for different values of h and k. The examples

$$h^2 + k^2, \quad -h^2 - k^2, \quad h^2 - k^2, \quad \pm(h-k)^2$$

show that a quadratic form may be: either 1°, positive for every (h, k) except $(0, 0)$; or 2°, negative for every (h, k) except $(0, 0)$; or 3°, positive for some values (h, k) and negative for others and zero for others; or finally 4°, zero for values other than $(0, 0)$, but either never negative or never positive. Moreover, the four possibilities here mentioned are the only cases conceivable except 5°, that $A = B = C = 0$ and the form always is 0. In the first case the form is called a *definite positive* form, in the second a *definite negative* form, in the third an *indefinite* form, and in the fourth and fifth a *singular* form. The first case assures a minimum, the second a maximum, the third neither a minimum nor a maximum (sometimes called a minimax); but the case of a singular form leaves the question entirely undecided just as the condition $f''(x) = 0$ did.

The conditions which distinguish between the different possibilities may be expressed in terms of the coefficients A, B, C.

1° pos. def.,	$B^2 < AC$,	$A, C > 0$;	3° indef.,	$B^2 > AC$;
2° neg. def.,	$B^2 < AC$,	$A, C < 0$;	4° sing.,	$B^2 = AC$.

The conditions for distinguishing between maxima and minima are:

$$\left.\begin{matrix} f'_x = 0 \\ f'_y = 0 \end{matrix}\right\} \; f''^2_{xy} < f''_{xx} f''_{yy}, \; \left\{\begin{matrix} f''_{xx}, f''_{yy} > 0 \text{ minimum;} \\ f''_{xx}, f''_{yy} < 0 \text{ maximum;} \end{matrix}\right. \tag{34}$$

$$f''^2_{xy} > f''_{xx} f''_{yy}, \text{ minimax}; \quad f''^2_{xy} = f''_{xx} f''_{yy}\ (?).$$

It may be noted that in applying these conditions to the case of a definite form it is sufficient to show that either f''_{xx} or f''_{yy} is positive or negative because they necessarily have the same sign.

EXERCISES

1. Write at length, without symbolic shortening, the expansion of $f(x, y)$ by Taylor's Formula to and including the terms of the third order in $x - a$, $y - b$. Write the formula also with the terms of the third order as the remainder.

2. Write by analogy the proper form of Taylor's Formula for $f(x, y, z)$ and prove it. Indicate the result for any number of variables.

3. Obtain the quadratic and lower terms in the development

(α) of $xy^2 + \sin xy$ at $(1, \frac{1}{2}\pi)$ and (β) of $\tan^{-1}(y/x)$ at $(1, 1)$.

4. A rectangular parallelepiped with one vertex at the origin and three faces in the coördinate planes has the opposite vertex upon the ellipsoid

$$x^2/a^2 + y^2/b^2 + z^2/c^2 = 1.$$

Find the maximum volume.

5. Find the point within a triangle such that the sum of the squares of its distances to the vertices shall be a minimum. Note that the point is the intersection of the medians. Is it obvious that a minimum and not a maximum is present?

6. A floating anchorage is to be made with a cylindrical body and equal conical ends. Find the dimensions that make the surface least for a given volume.

7. A cylindrical tent has a conical roof. Find the best dimensions.

8. Apply the test by second derivatives to the problem in the text and to any of Exs. 4–7. Discuss for maxima or minima the following functions:

(α) $x^2y + xy^2 - x$, (β) $x^3 + y^3 - x^2y^2 - \frac{1}{2}(x^2 + y^2)$,
(γ) $x^2 + y^2 + x + y$, (δ) $\frac{1}{3}y^3 - xy^2 + x^2y - x$,
(ϵ) $x^3 + y^3 - 9xy + 27$, (ζ) $x^4 + y^4 - 2x^2 + 4xy - 2y^2$.

9. State the conditions on the first derivatives for a maximum or minimum of function of three or any number of variables. Prove in the case of three variables.

10. A wall tent with rectangular body and gable roof is to be so constructed as to use the least amount of tenting for a given volume. Find the dimensions.

11. Given any number of masses $m_1, m_2, \cdots, m_n$ situated at (x_1, y_1), (x_2, y_2), $\cdots$, (x_n, y_n). Show that the point about which their moment of inertia is least is their center of gravity. If the points were (x_1, y_1, z_1), $\cdots$ in space, what point would make Σmr^2 a minimum?

12. A test for maximum or minimum analogous to that of Ex. 27, p. 10, may be given for a function $f(x, y)$ of two variables, namely: If a function is positive all over a region and vanishes upon the contour of the region, it must have a maximum within the region at the point for which $f'_x = f'_y = 0$. If a function is finite all over a region and becomes infinite over the contour of the region, it must have a minimum within the region at the point for which $f'_x = f'_y = 0$. These tests are subject to the proviso that $f'_x = f'_y = 0$ has only a single solution. Comment on the test and apply it to exercises above.

13. If a, b, c, r are the sides of a given triangle and the radius of the inscribed circle, the pyramid of altitude h constructed on the triangle as base will have its maximum surface when the surface is $\frac{1}{2}(a + b + c)\sqrt{r^2 + h^2}$.

CHAPTER V

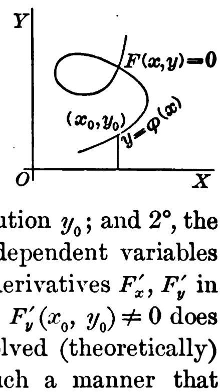
PARTIAL DIFFERENTIATION; IMPLICIT FUNCTIONS

56. The simplest case; $F(x, y) = 0$. The total differential

$$dF = F'_x dx + F'_y dy = d0 = 0$$

indicates
$$\frac{dy}{dx} = -\frac{F'_x}{F'_y}, \qquad \frac{dx}{dy} = -\frac{F'_y}{F'_x} \tag{1}$$

as the derivative of y by x, or of x by y, where y is defined as a function of x, or x as a function of y, by the relation $F(x, y) = 0$; and this method of obtaining a derivative of an *implicit function* without solving explicitly for the function has probably been familiar long before the notion of a partial derivative was obtained. The relation $F(x, y) = 0$ is pictured as a curve, and the function $y = \phi(x)$, which would be obtained by solution, is considered as multiple valued or as restricted to some definite portion or branch of the curve $F(x, y) = 0$. If the results (1) are to be applied to find the derivative at some point (x_0, y_0) of the curve $F(x, y) = 0$, it is necessary that at that point the denominator F'_y or F'_x should not vanish.

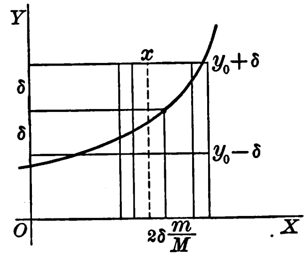


These pictorial and somewhat vague notions may be stated precisely as a *theorem* susceptible of proof, namely: Let x_0 be any real value of x such that 1°, the equation $F(x_0, y) = 0$ has a real solution y_0; and 2°, the function $F(x, y)$ regarded as a function of two independent variables (x, y) is continuous and has continuous first partial derivatives F'_x, F'_y in the neighborhood of (x_0, y_0); and 3°, the derivative $F'_y(x_0, y_0) \neq 0$ does not vanish for (x_0, y_0); then $F(x, y) = 0$ may be solved (theoretically) as $y = \phi(x)$ in the vicinity of $x = x_0$ and in such a manner that $y_0 = \phi(x_0)$, that $\phi(x)$ is continuous in x, and that $\phi(x)$ has a derivative $\phi'(x) = -F'_x/F'_y$; and the solution is unique. This is the fundamental theorem on implicit functions for the simple case, and the proof follows.

By the conditions on F'_x, F'_y, the Theorem of the Mean is applicable. Hence

$$F(x, y) - F(x_0, y_0) = F(x, y) = (hF'_x + kF'_y)_{x_0 + \theta h,\, y_0 + \theta k}. \tag{2}$$

Furthermore, in any square $|h| < \delta$, $|k| < \delta$ surrounding (x_0, y_0) and sufficiently small, the continuity of F'_x insures $|F'_x| < M$ and the continuity of F'_y taken with the fact that $F'_y(x_0, y_0) \neq 0$ insures $|F'_y| > m$. Consider the range of x as further restricted to values such that $|x - x_0| < m\delta/M$ if $m < M$. Now consider the value of $F(x, y)$ for any x in the permissible interval and for $y = y_0 + \delta$ or $y = y_0 - \delta$. As $|kF'_y| > m\delta$ but $|(x - x_0) F'_x| < m\delta$, it follows from (2) that $F(x, y_0 + \delta)$ has the sign of $\delta F'_y$ and $F(x, y_0 - \delta)$ has the sign of $-\delta F'_y$; and as the sign of F'_y does not change, $F(x, y_0 + \delta)$ and $F(x, y_0 - \delta)$ have opposite signs. Hence by Ex. 10, p. 45, there is one and only one value of y between $y_0 - \delta$ and $y_0 + \delta$ such that $F(x, y) = 0$. Thus for each x in the interval there is one and only one y such that $F(x, y) = 0$. The equation $F(x, y) = 0$ has a unique solution near (x_0, y_0). Let $y = \phi(x)$ denote the solution. The solution is continuous at $x = x_0$ because $|y - y_0| < \delta$. If (x, y) are restricted to values $y = \phi(x)$ such that $F(x, y) = 0$, equation (2) gives at once



$$\frac{k}{h} = \frac{y - y_0}{x - x_0} = \frac{\Delta y}{\Delta x} = -\frac{F'_x(x + \theta h, y + \theta k)}{F'_y(x + \theta h, y + \theta k)}, \qquad \frac{dy}{dx} = -\frac{F'_x(x_0, y_0)}{F'_y(x_0, y_0)}.$$

As F'_x, F'_y are continuous and $F'_y \neq 0$, the fraction k/h approaches a limit and the derivative $\phi'(x_0)$ exists and is given by (1). The same reasoning would apply to any point x in the interval. The theorem is completely proved. It may be added that the expression for $\phi'(x)$ is such as to show that $\phi'(x)$ itself is continuous.

The values of higher derivatives of implicit functions are obtainable by successive total differentiation as

$$F'_x + F'_y y' = 0,$$

$$F''_{xx} + 2 F''_{xy} y' + F''_{yy} y'^2 + F'_y y'' = 0, \tag{3}$$

etc. It is noteworthy that these successive equations may be solved for the derivative of highest order by dividing by F'_y which has been assumed not to vanish. The question of whether the function $y = \phi(x)$ defined implicitly by $F(x, y) = 0$ has derivatives of order higher than the first may be seen by these equations to depend on whether $F(x, y)$ has higher partial derivatives which are continuous in (x, y).

57. To find the *maxima and minima* of $y = \phi(x)$, that is, to find the points where the tangent to $F(x, y) = 0$ is parallel to the x-axis, observe that at such points $y' = 0$. Equations (3) give

$$F'_x = 0, \qquad F''_{xx} + F'_y y'' = 0. \tag{4}$$

Hence always under the assumption that $F'_y \neq 0$, *there are maxima at the intersections of $F = 0$ and $F'_x = 0$ if F''_{xx} and F'_y have the same sign, and minima at the intersections for which F''_{xx} and F'_y have opposite signs;* the case $F''_{xx} = 0$ still remains undecided.

For example if $F(x, y) = x^3 + y^3 - 3axy = 0$, the derivatives are

$$3(x^2 - ay) + 3(y^2 - ax)y' = 0, \qquad \frac{dy}{dx} = -\frac{x^2 - ay}{y^2 - ax},$$

$$6x - 6ay' + 6yy'^2 + 3(y^2 - ax)y'' = 0, \qquad \frac{d^2y}{dx^2} = -\frac{2a^3xy}{(y^2 - ax)^3}.$$

To find the maxima or minima of y as a function of x, solve

$$F'_x = 0 = x^2 - ay, \qquad F = 0 = x^3 + y^3 - 3axy, \qquad F'_y \neq 0.$$

The real solutions of $F'_x = 0$ and $F = 0$ are $(0, 0)$ and $(\sqrt[3]{2}\,a, \sqrt[3]{4}\,a)$ of which the first must be discarded because $F'_y(0, 0) = 0$. At $(\sqrt[3]{2}\,a, \sqrt[3]{4}\,a)$ the derivatives F'_y and F''_{xx} are positive; and the point is a maximum. The curve $F = 0$ is the folium of Descartes.

The rôle of the variables x and y may be interchanged if $F'_x \neq 0$ and the equation $F(x, y) = 0$ may be solved for $x = \psi(y)$, the functions ϕ and ψ being inverse. In this way the vertical tangents to the curve $F = 0$ may be discussed. For the points of $F = 0$ at which both $F'_x = 0$ and $F'_y = 0$, the equation cannot be solved in the sense here defined. Such points are called *singular points* of the curve. The questions of the singular points of $F = 0$ and of maxima, minima, or minimax (§ 55) of the surface $z = F(x, y)$ are related. For if $F'_x = F'_y = 0$, the surface has a tangent plane parallel to $z = 0$, and if the condition $z = F = 0$ is also satisfied, the surface is tangent to the xy-plane. Now if $z = F(x, y)$ has a maximum or minimum at its point of tangency with $z = 0$, the surface lies entirely on one side of the plane and the point of tangency is an isolated point of $F(x, y) = 0$; whereas if the surface has a minimax it cuts through the plane $z = 0$ and the point of tangency is not an isolated point of $F(x, y) = 0$. The shape of the curve $F = 0$ in the neighborhood of a singular point is discussed by developing $F(x, y)$ about that point by Taylor's Formula.

For example, consider the curve $F(x, y) = x^3 + y^3 - x^2y^2 - \frac{1}{2}(x^2 + y^2) = 0$ and the surface $z = F(x, y)$. The common real solutions of

$$F'_x = 3x^2 - 2xy^2 - x = 0, \qquad F'_y = 3y^2 - 2x^2y - y = 0, \qquad F(x, y) = 0$$

are the singular points. The real solutions of $F'_x = 0$, $F'_y = 0$ are $(0, 0)$, $(1, 1)$, $(\frac{1}{3}, \frac{1}{3})$ and of these the first two satisfy $F(x, y) = 0$ but the last does not. The singular points of the curve are therefore $(0, 0)$ and $(1, 1)$. The test (34) of § 55 shows that $(0, 0)$ is a maximum for $z = F(x, y)$ and hence an isolated point of $F(x, y) = 0$. The test also shows that $(1, 1)$ is a minimax. To discuss the curve $F(x, y) = 0$ near $(1, 1)$ apply Taylor's Formula.

$$\begin{aligned} 0 = F(x, y) &= \tfrac{1}{2}(3h^2 - 8hk + 3k^2) + \tfrac{1}{6}(6h^3 - 12h^2k - 12hk^2 + 6k^3) + \text{remainder} \\ &= \tfrac{1}{2}(3\cos^2\phi - 8\sin\phi\cos\phi + 3\sin^2\phi) \\ &\qquad + r(\cos^3\phi - 2\cos^2\phi\sin\phi - 2\cos\phi\sin^2\phi + \sin^3\phi) + \cdots, \end{aligned}$$

if polar coördinates $h = r\cos\phi$, $k = r\sin\phi$ be introduced at (1, 1) and r^2 be canceled. Now for very small values of r, the equation can be satisfied only when the first parenthesis is very small. Hence the solutions of

$$3 - 4\sin 2\phi = 0, \qquad \sin 2\phi = \tfrac{3}{4}, \qquad \text{or} \qquad \phi = 24^\circ\, 17\tfrac{1}{2}',\ 65^\circ\, 42\tfrac{1}{2}',$$

and $\phi + \pi$, are the directions of the tangents to $F(x, y) = 0$. The equation $F = 0$ is

$$0 = (1\tfrac{1}{2} - 2\sin 2\phi) + r(\cos\phi + \sin\phi)(1 - 1\tfrac{1}{2}\sin 2\phi)$$

if only the first two terms are kept, and this will serve to sketch $F(x, y) = 0$ for very small values of r, that is, for ϕ very near to the tangent directions.

58. It is important to obtain conditions for the maximum or minimum of a function $z = f(x, y)$ where the variables x, y are connected by a relation $F(x, y) = 0$ so that z really becomes a function of x alone or y alone. For it is not always possible, and frequently it is inconvenient, to solve $F(x, y) = 0$ for either variable and thus eliminate that variable from $z = f(x, y)$ by substitution. When the variables x, y in $z = f(x, y)$ are thus connected, the minimum or maximum is called a *constrained minimum* or *maximum;* when there is no equation $F(x, y) = 0$ between them the minimum or maximum is called *free* if any designation is needed.* The conditions are obtained by differentiating $z = f(x, y)$ and $F(x, y) = 0$ totally with respect to x. Thus

$$\frac{dz}{dx} = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}\frac{dy}{dx} = 0, \qquad \frac{d0}{dx} = \frac{\partial F}{\partial x} + \frac{\partial F}{\partial y}\frac{dy}{dx} = 0,$$

and

$$\frac{\partial f}{\partial x}\frac{\partial F}{\partial y} - \frac{\partial f}{\partial y}\frac{\partial F}{\partial x} = 0, \qquad \frac{d^2z}{dx^2} \gtrless 0, \qquad F = 0, \tag{5}$$

where the first equation arises from the two above by eliminating dy/dx and the second is added to insure a minimum or maximum, are the conditions desired. Note that all singular points of $F(x, y) = 0$ satisfy the first condition identically, but that the process by means of which it was obtained excludes such points, and that the rule cannot be expected to apply to them.

Another method of treating the problem of constrained maxima and minima is to introduce *a multiplier* and form the function

$$z = \Phi(x, y) = f(x, y) + \lambda F(x, y), \qquad \lambda \text{ a multiplier.} \tag{6}$$

Now if this function z is to have a free maximum or minimum, then

$$\Phi'_x = f'_x + \lambda F'_x = 0, \qquad \Phi'_y = f'_y + \lambda F'_y = 0. \tag{7}$$

These two equations taken with $F = 0$ constitute a set of three from which the three values x, y, λ may be obtained by solution. Note that

* The adjective "relative" is sometimes used for constrained, and "absolute" for free; but the term "absolute" is best kept for the greatest of the maxima or least of the minima, and the term "relative" for the other maxima and minima.

λ cannot be obtained from (7) if both F'_x and F'_y vanish; and hence this method also rejects the singular points. That this method really determines the constrained maxima and minima of $f(x, y)$ subject to the constraint $F(x, y) = 0$ is seen from the fact that if λ be eliminated from (7) the condition $f'_x F'_y - f'_y F'_x = 0$ of (5) is obtained. The new method is therefore identical with the former, and its introduction is more a matter of convenience than necessity. It is possible to show directly that the new method gives the constrained maxima and minima. For the conditions (7) are those of a free extreme for the function $\Phi(x, y)$ which depends on two independent variables (x, y). Now if the equations (7) be solved for (x, y), it appears that the position of the maximum or minimum will be expressed in terms of λ as a parameter and that consequently the point $(x(\lambda), y(\lambda))$ cannot in general lie on the curve $F(x, y) = 0$; but if λ be so determined that the point shall lie on this curve, the function $\Phi(x, y)$ has a free extreme at a point for which $F = 0$ and hence in particular must have a constrained extreme for the particular values for which $F(x, y) = 0$. In speaking of (7) as the conditions for an extreme, the conditions which should be imposed on the second derivative have been disregarded.

For example, suppose the maximum radius vector from the origin to the folium of Descartes were desired. The problem is to render $f(x, y) = x^2 + y^2$ maximum subject to the condition $F(x, y) = x^3 + y^3 - 3axy = 0$. Hence

$$2x + 3\lambda(x^2 - ay) = 0, \qquad 2y + 3\lambda(y^2 - ax) = 0, \qquad x^3 + y^3 - 3axy = 0$$

or

$$2x \cdot 3(y^2 - ax) - 2y \cdot 3(x^2 - ay) = 0, \qquad x^3 + y^3 - 3axy = 0$$

are the conditions in the two cases. These equations may be solved for (0, 0), $(1\frac{1}{2}a, 1\frac{1}{2}a)$, and some imaginary values. The value (0, 0) is singular and λ cannot be determined, but the point is evidently a minimum of $x^2 + y^2$ by inspection. The point $(1\frac{1}{2}a, 1\frac{1}{2}a)$ gives $\lambda = -1\frac{1}{3}a$. That the point is a (relative constrained) maximum of $x^2 + y^2$ is also seen by inspection. There is no need to examine d^2f. In most practical problems the examination of the conditions of the second order may be waived. This example is one which may be treated in polar coördinates by the ordinary methods; but it is noteworthy that if it could not be treated that way, the method of solution by eliminating one of the variables by solving the cubic $F(x, y) = 0$ would be unavailable and the methods of constrained maxima would be required.

EXERCISES

1. By total differentiation and division obtain dy/dx in these cases. Do not substitute in (1), but use the method by which it was derived.

(α) $ax^2 + 2bxy + cy^2 - 1 = 0$, (β) $x^4 + y^4 = 4a^2xy$, (γ) $(\cos x)^y - (\sin y)^x = 0$,
(δ) $(x^2 + y^2)^2 = a^2(x^2 - y^2)$, (ϵ) $e^x + e^y = 2xy$, (ζ) $x^{-2}y^{-\frac{1}{2}} = \tan^{-1} xy$.

2. Obtain the second derivative d^2y/dx^2 in Ex. 1 (α), (β), (ϵ), (ζ) by differentiating the value of dy/dx obtained above. Compare with use of (3).

3. Prove $\dfrac{d^2y}{dx^2} = -\dfrac{F_y'^2F_{xx}'' - 2F_x'F_y'F_{xy}'' + F_x'^2F_{yy}''}{F_y'^3}$.

4. Find the radius of curvature of these curves:

(α) $x^{\frac{2}{3}} + y^{\frac{2}{3}} = a^{\frac{2}{3}}$, $R = 3(axy)^{\frac{1}{3}}$, (β) $x^{\frac{1}{2}} + y^{\frac{1}{2}} = a^{\frac{1}{2}}$, $R = 2\sqrt{(x+y)^3/a}$,
(γ) $b^2x^2 + a^2y^2 = a^2b^2$, (δ) $xy^2 = a^2(a-x)$, (ϵ) $(ax)^2 + (by)^{\frac{2}{3}} = 1$.

5. Find y', y'', y''' in case $x^3 + y^3 - 3axy = 0$.

6. Extend equations (3) to obtain y''' and reduce by Ex. 3.

7. Find tangents parallel to the x-axis for $(x^2+y^2)^2 = 2a^2(x^2-y^2)$.

8. Find tangents parallel to the y-axis for $(x^2+y^2+ax)^2 = a^2(x^2+y^2)$.

9. If $b^2 < ac$ in $ax^2 + 2bxy + cy^2 + fx + gy + h = 0$, circumscribe about the curve a rectangle parallel to the axes. Check algebraically.

10. Sketch $x^3 + y^3 = x^2y^2 + \frac{1}{2}(x^2+y^2)$ near the singular point (1, 1).

11. Find the singular points and discuss the curves near them:

(α) $x^3 + y^3 = 3axy$, (β) $(x^2+y^2)^2 = 2a^2(x^2-y^2)$,
(γ) $x^4 + y^4 = 2(x-y)^2$, (δ) $y^5 + 2xy^2 = x^2 + y^4$.

12. Make these functions maxima or minima subject to the given conditions. Discuss the work both with and without a multiplier:

(α) $\dfrac{a}{u\cos x} + \dfrac{b}{v\cos y}$, $a\tan x + b\tan y = c$. *Ans.* $\dfrac{\sin x}{\sin y} = \dfrac{u}{v}$.
(β) $x^2 + y^2$, $ax^2 + 2bxy + cy^2 = f$. Find axes of conic.
(γ) Find the shortest distance from a point to a line (in a plane).

13. Write the second and third total differentials of $F(x, y) = 0$ and compare with (3) and Ex. 5. Try this method of calculating in Ex. 2.

14. Show that $F_x'dx + F_y'dy = 0$ does and should give the tangent line to $F(x, y) = 0$ at the points (x, y) if $dx = \xi - x$ and $dy = \eta - y$, where ξ, η are the coördinates of points other than (x, y) on the tangent line. Why is the equation inapplicable at singular points of the curve?

59. More general cases of implicit functions. The problem of implicit functions may be generalized in two ways. In the first place a greater number of variables may occur in the function, as

$$F(x, y, z) = 0, \qquad F(x, y, z, \cdots, u) = 0;$$

and the question may be to solve the equation for one of the variables in terms of the others and to determine the partial derivatives of the chosen dependent variable. In the second place there may be several equations connecting the variables and it may be required to solve the equations for some of the variables in terms of the others and to determine the partial derivatives of the chosen dependent variables with respect to the independent variables. In both cases the formal differentiation and attempted formal solution of the equations for the derivatives will indicate the results and the theorem under which the solution is proper.

Consider the case $F(x, y, z) = 0$ and form the differential.

$$dF(x, y, z) = F'_x dx + F'_y dy + F'_z dz = 0. \tag{8}$$

If z is to be the dependent variable, the partial derivative of z by x is found by setting $dy = 0$ so that y is constant. Thus

$$\frac{\partial z}{\partial x} = \left(\frac{dz}{dx}\right)_y = -\frac{F'_x}{F'_z} \quad \text{and} \quad \frac{\partial z}{\partial y} = \left(\frac{dz}{dy}\right)_x = -\frac{F'_y}{F'_z} \tag{9}$$

are obtained by ordinary division after setting $dy = 0$ and $dx = 0$ respectively. If this division is to be legitimate, F'_z must not vanish at the point considered. The immediate suggestion is the theorem: If, when real values (x_0, y_0) are chosen and a real value z_0 is obtained from $F(z, x_0, y_0) = 0$ by solution, the function $F(x, y, z)$ regarded as a function of three independent variables (x, y, z) is continuous at and near (x_0, y_0, z_0) and has continuous first partial derivatives and $F'_z(x_0, y_0, z_0) \neq 0$, then $F(x, y, z) = 0$ may be solved uniquely for $z = \phi(x, y)$ and $\phi(x, y)$ will be continuous and have partial derivatives (9) for values of (x, y) sufficiently near to (x_0, y_0).

The theorem is again proved by the Law of the Mean, and in a similar manner.

$$F(x, y, z) - F(x_0, y_0, z_0) = F(x, y, z) = (hF'_x + kF'_y + lF'_z)_{x_0+\theta h,\, y_0+\theta k,\, z_0+\theta l}.$$

As F'_x, F'_y, F'_z are continuous and $F'_z(x_0, y_0, z_0) \neq 0$, it is possible to take δ so small that, when $|h| < \delta$, $|k| < \delta$, $|l| < \delta$, the derivative $|F'_z| > m$ and $|F'_x| < \mu$, $|F'_y| < \mu$. Now it is desired so to restrict h, k that $\pm \delta F'_z$ shall determine the sign of the parenthesis. Let

$$|x - x_0| < \tfrac{1}{2} m\delta/\mu, \qquad |y - y_0| < \tfrac{1}{2} m\delta/\mu, \quad \text{then} \quad |hF'_x + kF'_y| < m\delta$$

and the signs of the parenthesis for $(x, y, z_0 + \delta)$ and $(x, y, z_0 - \delta)$ will be opposite since $|F'_z| > m$. Hence if (x, y) be held fixed, there is one and only one value of z for which the parenthesis vanishes between $z_0 + \delta$ and $z_0 - \delta$. Thus z is defined as a single valued function of (x, y) for sufficiently small values of $h = x - x_0$, $k = y - y_0$.

Also $$\frac{l}{h} = -\frac{F'_x(x_0 + \theta h, y_0 + \theta k, z_0 + \theta l)}{F'_z(x_0 + \theta h, y_0 + \theta k, z_0 + \theta l)}, \qquad \frac{l}{k} = -\frac{F'_y(\cdots)}{F'_z(\cdots)}$$

when k and h respectively are assigned the values 0. The limits exist when $h \doteq 0$ or $k \doteq 0$. But in the first case $l = \Delta z = \Delta_x z$ is the increment of z when x alone varies, and in the second case $l = \Delta z = \Delta_y z$. The limits are therefore the desired partial derivatives of z by x and y. The proof for any number of variables would be similar.

If none of the derivatives F'_x, F'_y, F'_z vanish, the equation $F(x, y, z) = 0$ may be solved for any one of the variables, and formulas like (9) will express the partial derivatives. It then appears that

$$\left(\frac{dz}{dx}\right)_y\left(\frac{dx}{dz}\right)_y = \frac{\partial z}{\partial x}\frac{\partial x}{\partial z} = \frac{F'_x}{F'_z}\frac{F'_z}{F'_x} = 1, \tag{10}$$

and

$$\left(\frac{dz}{dx}\right)_y\left(\frac{dx}{dy}\right)_z\left(\frac{dy}{dz}\right)_x = \frac{\partial z}{\partial x}\frac{\partial x}{\partial y}\frac{\partial y}{\partial z} = -1 \tag{11}$$

in like manner. The first equation is in this case identical with (4) of § 2 because if y is constant the relation $F(x, y, z) = 0$ reduces to $G(x, z) = 0$. The second equation is new. By virtue of (10) and similar relations, the derivatives in (11) may be inverted and transformed to the right side of the equation. As it is assumed in thermodynamics that the pressure, volume, and temperature of a given simple substance are connected by an equation $F(p, v, T) = 0$, called the characteristic equation of the substance, a relation between different thermodynamic magnitudes is furnished by (11).

60. In the next place suppose there are two equations

$$F(x, y, u, v) = 0, \qquad G(x, y, u, v) = 0 \tag{12}$$

between four variables. Let each equation be differentiated.

$$dF = 0 = F'_x dx + F'_y dy + F'_u du + F'_v dv,$$
$$dG = 0 = G'_x dx + G'_y dy + G'_u du + G'_v dv. \tag{13}$$

If it be desired to consider u, v as the dependent variables and x, y as independent, it would be natural to solve these equations for the differentials du and dv in terms of dx and dy; for example,

$$du = \frac{(F'_x G'_v - F'_v G'_x)\,dx + (F'_y G'_v - F'_v G'_y)\,dy}{F'_u G'_v - F'_v G'_u}. \tag{13'}$$

The differential dv would have a different numerator but the same denominator. The solution requires $F'_u G'_v - F'_v G'_u \neq 0$. This suggests the desired theorem: If (u_0, v_0) are solutions of $F = 0$, $G = 0$ corresponding to (x_0, y_0) and if $F'_u G'_v - F'_v G'_u$ does not vanish for the values (x_0, y_0, u_0, v_0), the equations $F = 0$, $G = 0$ may be solved for $u = \phi(x, y)$, $v = \psi(x, y)$ and the solution is unique and valid for (x, y) sufficiently near (x_0, y_0) — it being assumed that F and G regarded as functions in four variables are continuous and have continuous first partial derivatives at and near (x_0, y_0, u_0, v_0); moreover, the total differentials du, dv are given by (13′) and a similar equation.

The proof of this theorem may be deferred (§ 64). Some observations should be made. The equations (13) may be solved for any two variables in terms of the other two. The partial derivatives

$$\frac{\partial u(x, y)}{\partial x}, \quad \frac{\partial u(x, v)}{\partial x}, \quad \frac{\partial x(u, v)}{\partial u}, \quad \frac{\partial x(u, y)}{\partial u} \tag{14}$$

of u by x or of x by u will naturally depend on whether the solution for u is in terms of (x, y) or of (x, v), and the solution for x is in (u, v) or (u, y). Moreover, it must not be assumed that $\partial u/\partial x$ and $\partial x/\partial u$ are reciprocals no matter which meaning is attached to each. In obtaining relations between the derivatives analogous to (10), (11), the values of the derivatives in terms of the derivatives of F and G may be found or the equations (12) may first be considered as solved.

Thus if
$$u = \phi(x, y), \qquad du = \phi_x' dx + \phi_y' dy,$$
$$v = \psi(x, y), \qquad dv = \psi_x' dx + \psi_y' dy.$$

Then
$$dx = \frac{\psi_y' du - \phi_y' dv}{\phi_x'\psi_y' - \phi_y'\psi_x'}, \qquad dy = \frac{-\psi_x' du + \phi_x' dv}{\phi_x'\psi_y' - \phi_y'\psi_x'}$$

and
$$\frac{\partial x}{\partial u} = \frac{\psi_y'}{\phi_x'\psi_y' - \phi_y'\psi_x'}, \qquad \frac{\partial x}{\partial v} = \frac{-\phi_y'}{\phi_x'\psi_y' - \phi_y'\psi_x'}, \text{ etc.}$$

Hence
$$\frac{\partial u}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial v}{\partial x}\frac{\partial x}{\partial v} = 1, \tag{15}$$

as may be seen by direct substitution. Here u, v are expressed in terms of x, y for the derivatives u_x', v_x'; and x, y are considered as expressed in terms of u, v for the derivatives x_u', x_v'.

61. The questions of free or constrained maxima and minima, at any rate in so far as the determination of the conditions of the first order is concerned, may now be treated. If $F(x, y, z) = 0$ is given and the maxima and minima of z as a function of (x, y) are wanted,

$$F_x'(x, y, z) = 0, \qquad F_y'(x, y, z) = 0, \qquad F(x, y, z) = 0 \tag{16}$$

are three equations which may be solved for x, y, z. If for any of these solutions the derivative F_z' does not vanish, the surface $z = \phi(x, y)$ has at that point a tangent plane parallel to $z = 0$ and there is a maximum, minimum, or minimax. To distinguish between the possibilities further investigation must be made if necessary; the details of such an investigation will not be outlined for the reason that special methods are usually available. The conditions for an extreme of u as a function of (x, y) defined implicitly by the equations (13′) are seen to be

$$F_x'G_v' - F_v'G_x' = 0, \quad F_y'G_v' - F_v'G_y' = 0, \quad F = 0, \quad G = 0. \tag{17}$$

The four equations may be solved for x, y, u, v or merely for x, y.

Suppose that the maxima, minima, and minimax of $u = f(x, y, z)$ subject either to one equation $F(x, y, z) = 0$ or two equations $F(x, y, z) = 0$, $G(x, y, z) = 0$ of constraint are desired. Note that if only one equation of constraint is imposed, the function $u = f(x, y, z)$ becomes a function of two variables; whereas if two equations are imposed, the function u really contains only one variable and the question of a minimax does not arise. The *method of multipliers* is again employed. Consider

$$\Phi(x, y, z) = f + \lambda F \quad \text{or} \quad \Phi = f + \lambda F + \mu G \tag{18}$$

as the case may be. The conditions for a free extreme of Φ are

$$\Phi'_x = 0, \qquad \Phi'_y = 0, \qquad \Phi'_z = 0. \tag{19}$$

These three equations may be solved for the coördinates x, y, z which will then be expressed as functions of λ or of λ and μ according to the case. If then λ or λ and μ be determined so that (x, y, z) satisfy $F = 0$ or $F = 0$ and $G = 0$, the constrained extremes of $u = f(x, y, z)$ will be found except for the examination of the conditions of higher order.

As a problem in constrained maxima and minima let the axes of the section of an ellipsoid by a plane through the origin be determined. Form the function

$$\Phi = x^2 + y^2 + z^2 + \lambda\left(\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} - 1\right) + \mu(lx + my + nz)$$

by adding to $x^2 + y^2 + z^2$, which is to be made extreme, the equations of the ellipsoid and plane, which are the equations of constraint. Then apply (19). Hence

$$x + \lambda\frac{x}{a^2} + \frac{\mu}{2}l = 0, \qquad y + \lambda\frac{y}{b^2} + \frac{\mu}{2}m = 0, \qquad z + \lambda\frac{z}{c^2} + \frac{\mu}{2}n = 0$$

taken with the equations of ellipsoid and plane will determine x, y, z, λ, μ. If the equations are multiplied by x, y, z and reduced by the equations of plane and ellipsoid, the solution for λ is $\lambda = -r^2 = -(x^2 + y^2 + z^2)$. The three equations then become

$$x = \frac{1}{2}\frac{\mu l a^2}{r^2 - a^2}, \qquad y = \frac{1}{2}\frac{\mu m b^2}{r^2 - b^2}, \qquad z = \frac{1}{2}\frac{\mu n c^2}{r^2 - c^2}, \qquad \text{with } lx + my + nz = 0.$$

Hence
$$\frac{l^2a^2}{r^2 - a^2} + \frac{m^2b^2}{r^2 - b^2} + \frac{n^2c^2}{r^2 - c^2} = 0 \quad \text{determines } r^2. \tag{20}$$

The two roots for r are the major and minor axes of the ellipse in which the plane cuts the ellipsoid. The substitution of x, y, z above in the ellipsoid determines

$$\frac{\mu^2}{4} = \left(\frac{al}{r^2 - a^2}\right)^2 + \left(\frac{bm}{r^2 - b^2}\right)^2 + \left(\frac{cn}{r^2 - c^2}\right)^2 \quad \text{since} \quad \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1. \tag{21}$$

Now when (20) is solved for any particular root r and the value of μ is found by (21), the actual coördinates x, y, z of the extremities of the axes may be found.

EXERCISES

1. Obtain the partial derivatives of z by x and y directly from (8) and not by substitution in (9). Where does the solution fail?

(α) $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$, $\quad(\beta)$ $x + y + z = \frac{1}{xyz}$,

(γ) $(x^2 + y^2 + z^2)^2 = a^2x^2 + b^2y^2 + c^2z^2$, $\quad(\delta)$ $xyz = c$.

2. Find the second derivatives in Ex. 1 (α), (β), (δ) by repeated differentiation.

3. State and prove the theorem on the solution of $F(x, y, z, u) = 0$.

4. Show that the product $\alpha_p E_T$ of the coefficient of expansion by the modulus of elasticity (§ 52) is equal to the rate of rise of pressure with the temperature if the volume is constant.

5. Establish the proportion $E_S : E_T = C_p : C_v$ (see § 52).

6. If $F(x, y, z, u) = 0$, show $\frac{\partial u}{\partial x}\frac{\partial x}{\partial y}\frac{\partial y}{\partial z}\frac{\partial z}{\partial u} = 1$, $\frac{\partial u}{\partial x}\frac{\partial x}{\partial u} = 1$.

7. Write the equations of tangent plane and normal line to $F(x, y, z) = 0$ and find the tangent planes and normal lines to Ex. 1 (β), (δ) at $x = 1$, $y = 1$.

8. Find, by using (13), the indicated derivatives on the assumption that either x, y or u, v are dependent and the other pair independent:

(α) $u^5 + v^5 + x^5 - 3y = 0$, $\quad u^3 + v^3 + y^3 + 3x = 0$, $\quad u'_x, u'_y, u''_{xy}, v''_{xx}$

(β) $x + y + u + v = a$, $\quad x^2 + y^2 + u^2 + v^2 = b$, $\quad x'_u, u'_x, v'_y, v''_{xy}$

(γ) Find dy in both cases if x, v are independent variables.

9. Prove $\frac{\partial u}{\partial x}\frac{\partial y}{\partial u} + \frac{\partial v}{\partial x}\frac{\partial y}{\partial v} = 0$ if $F(x, y, u, v) = 0$, $G(x, y, u, v) = 0$.

10. Find du and the derivatives u'_x, u'_y, u'_z in case

$$x^2 + y^2 + z^2 = uv, \qquad xy = u^2 + v^2 + w^2, \qquad xyz = uvw.$$

11. If $F(x, y, z) = 0$, $G(x, y, z) = 0$ define a curve, show that

$$\frac{x - x_0}{(F'_yG'_z - F'_zG'_y)_0} = \frac{y - y_0}{(F'_zG'_x - F'_xG'_z)_0} = \frac{z - z_0}{(F'_xG'_y - F'_yG'_x)_0}$$

is the tangent line to the curve at (x_0, y_0, z_0). Write the normal plane.

12. Formulate the problem of implicit functions occurring in Ex. 11.

13. Find the perpendicular distance from a point to a plane.

14. The sum of three positive numbers is $x + y + z = N$, where N is given. Determine x, y, z so that the product $x^py^qz^r$ shall be maximum if p, q, r are given.

Ans. $x : y : z : N = p : q : r : (p + q + r)$.

15. The sum of three positive numbers and the sum of their squares are both given. Make the product a maximum or minimum.

16. The surface $(x^2+y^2+z^2)^2=ax^2+by^2+cz^2$ is cut by the plane $lx+my+nz=0$. Find the maximum or minimum radius of the section.

Ans. $\sum \frac{l^2}{r^2 - a} = 0$.

17. In case $F(x, y, u, v) = 0$, $G(x, y, u, v) = 0$ consider the differentials

$$dv = \frac{\partial v}{\partial x}dx + \frac{\partial v}{\partial y}dy, \qquad dx = \frac{\partial x}{\partial u}du + \frac{\partial x}{\partial v}dv, \qquad dy = \frac{\partial y}{\partial u}du + \frac{\partial y}{\partial v}dv.$$

Substitute in the first from the last two and obtain relations like (15) and Ex. 9.

18. If $f(x, y, z)$ is to be maximum or minimum subject to the constraint $F(x, y, z) = 0$, show that the conditions are that $dx : dy : dz = 0 : 0 : 0$ are indeterminate when their solution is attempted from

$$f'_x dx + f'_y dy + f'_z dz = 0 \quad \text{and} \quad F'_x dx + F'_y dy + F'_z dz = 0.$$

From what geometrical considerations should this be obvious? Discuss in connection with the problem of inscribing the maximum rectangular parallelepiped in the ellipsoid. These equations,

$$dx : dy : dz = f'_y F'_z - f'_z F'_y : f'_z F'_x - f'_x F'_z : f'_x F'_y - f'_y F'_x = 0 : 0 : 0,$$

may sometimes be used to advantage for such problems.

19. Given the curve $F(x, y, z) = 0$, $G(x, y, z) = 0$. Discuss the conditions for the highest or lowest points, or more generally the points where the tangent is parallel to $z = 0$, by treating $u = f(x, y, z) = z$ as a maximum or minimum subject to the two constraining equations $F = 0$, $G = 0$. Show that the condition $F'_x G'_y = F'_y G'_x$ which is thus obtained is equivalent to setting $dz = 0$ in

$$F'_x dx + F'_y dy + F'_z dz = 0 \quad \text{and} \quad G'_x dx + G'_y dy + G'_z dz = 0.$$

20. Find the highest and lowest points of these curves:

(α) $x^2 + y^2 = z^2 + 1$, $x + y + 2z = 0$, $\qquad (\beta)$ $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} + \dfrac{z^2}{c^2} = 1$, $lx + my + nz = 0$.

21. Show that $F'_x dx + F'_y dy + F'_z dz = 0$, with $dx = \xi - x$, $dy = \eta - y$, $dz = \zeta - z$, is the tangent plane to the surface $F(x, y, z) = 0$ at (x, y, z). Apply to Ex. 1.

22. Given $F(x, y, u, v) = 0$, $G(x, y, u, v) = 0$. Obtain the equations

$$\frac{\partial F}{\partial x} + \frac{\partial F}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial F}{\partial v}\frac{\partial v}{\partial x} = 0, \qquad \frac{\partial F}{\partial y} + \frac{\partial F}{\partial u}\frac{\partial u}{\partial y} + \frac{\partial F}{\partial v}\frac{\partial v}{\partial y} = 0,$$

$$\frac{\partial G}{\partial x} + \frac{\partial G}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial G}{\partial v}\frac{\partial v}{\partial x} = 0, \qquad \frac{\partial G}{\partial y} + \frac{\partial G}{\partial u}\frac{\partial u}{\partial y} + \frac{\partial G}{\partial v}\frac{\partial v}{\partial y} = 0,$$

and explain their significance as a sort of partial-total differentiation of $F = 0$ and $G = 0$. Find u'_x from them and compare with (13′). Write similar equations where x, y are considered as functions of (u, v). Hence prove, and compare with (15) and Ex. 9,

$$\frac{\partial u}{\partial y}\frac{\partial y}{\partial u} + \frac{\partial v}{\partial y}\frac{\partial y}{\partial v} = 1, \qquad \frac{\partial u}{\partial y}\frac{\partial x}{\partial u} + \frac{\partial v}{\partial y}\frac{\partial x}{\partial v} = 0.$$

23. Show that the differentiation with respect to x and y of the four equations under Ex. 22 leads to eight equations from which the eight derivatives

$$\frac{\partial^2 u}{\partial x^2}, \quad \frac{\partial^2 u}{\partial x \partial y}, \quad \frac{\partial^2 u}{\partial y \partial x}, \quad \frac{\partial^2 u}{\partial y^2}, \quad \frac{\partial^2 v}{\partial x^2}, \quad \cdots, \quad \frac{\partial^2 v}{\partial y^2}$$

may be obtained. Show thus that formally $u''_{xy} = u''_{yx}$.

62. Functional determinants or Jacobians. Let two functions

$$u = \phi(x, y), \qquad v = \psi(x, y) \tag{22}$$

of two independent variables be given. The continuity of the functions and of their first derivatives is assumed throughout this discussion and will not be mentioned again. Suppose that there were a relation $F(u, v) = 0$ or $F(\phi, \psi) = 0$ between the functions. Then

$$F(\phi, \psi) = 0, \qquad F'_u\phi'_x + F'_v\psi'_x = 0, \qquad F'_u\phi'_y + F'_v\psi'_y = 0. \tag{23}$$

The last two equations arise on differentiating the first with respect to x and y. The elimination of F'_u and F'_v from these gives

$$\phi'_x\psi'_y - \phi'_y\psi'_x = \begin{vmatrix} \phi'_x & \psi'_x \\ \phi'_y & \psi'_y \end{vmatrix} = \frac{\partial(u, v)}{\partial(x, y)} = J\left(\frac{u, v}{x, y}\right) = 0. \tag{24}$$

The determinant is merely another way of writing the first expression; the next form is the customary short way of writing the determinant and denotes that the elements of the determinant are the first derivatives of u and v with respect to x and y. This determinant is called the *functional determinant* or *Jacobian* of the functions u, v or ϕ, ψ with respect to the variables x, y and is denoted by J. It is seen that: *If there is a functional relation $F(\phi, \psi) = 0$ between two functions, the Jacobian of the functions vanishes identically*, that is, vanishes for all values of the variables (x, y) under consideration.

Conversely, *if the Jacobian vanishes identically over a two-dimensional region for (x, y), the functions are connected by a functional relation.* For, the functions u, v may be assumed not to reduce to mere constants and hence there may be assumed to be points for which at least one of the partial derivatives ϕ'_x, ϕ'_y, ψ'_x, ψ'_y does not vanish. Let ϕ'_x be the derivative which does not vanish at some particular point of the region. Then $u = \phi(x, y)$ may be solved as $x = \chi(u, y)$ in the vicinity of that point and the result may be substituted in v.

$$v = \psi(\chi, y), \qquad \frac{\partial v}{\partial y} = \psi'_x\frac{\partial \chi}{\partial y} + \psi'_y = \psi'_x\frac{\partial x}{\partial y} + \psi'_y.$$

But
$$\frac{\partial x}{\partial y} = -\frac{\partial u}{\partial y}\frac{\partial x}{\partial u} \quad \text{and} \quad \frac{\partial v}{\partial y} = \frac{1}{\phi'_x}(\phi'_x\psi'_y - \psi'_x\phi'_y) \tag{24'}$$

by (11) and substitution. Thus $\partial v/\partial y = J/\phi'_x$; and if $J = 0$, then $\partial v/\partial y = 0$. This relation holds at least throughout the region for which $\phi'_x \neq 0$, and for points in this region $\partial v/\partial y$ vanishes identically. Hence v does not depend on y but becomes a function of u alone. This establishes the fact that v and u are functionally connected.

These considerations may be extended to other cases. Let

$$u = \phi(x, y, z), \qquad v = \psi(x, y, z), \qquad w = \chi(x, y, z). \tag{25}$$

If there is a functional relation $F(u, v, w) = 0$, differentiate it.

$$\begin{aligned} F'_u\phi'_x + F'_v\psi'_x + F'_w\chi'_x &= 0, \\ F'_u\phi'_y + F'_v\psi'_y + F'_w\chi'_y &= 0, \\ F'_u\phi'_z + F'_v\psi'_z + F'_w\chi'_z &= 0, \end{aligned} \qquad \begin{vmatrix} \phi'_x & \psi'_x & \chi'_x \\ \phi'_y & \psi'_y & \chi'_y \\ \phi'_z & \psi'_z & \chi'_z \end{vmatrix} = 0, \tag{26}$$

or

$$\frac{\partial(\phi, \psi, \chi)}{\partial(x, y, z)} = \frac{\partial(u, v, w)}{\partial(x, y, z)} = J = 0.$$

The result is obtained by eliminating F'_u, F'_v, F'_w from the three equations. The assumption is made, here as above, that F'_u, F'_v, F'_w do not all vanish; for if they did, the three equations would not imply $J = 0$. On the other hand their vanishing would imply that F did not contain u, v, w, — as it must if there is really a relation between them. And now conversely it may be shown that if J vanishes identically, there is a functional relation between u, v, w. Hence again *the necessary and sufficient conditions that the three functions* (25) *be functionally connected is that their Jacobian vanish.*

The proof of the converse part is about as before. It may be assumed that at least one of the derivatives of u, v, w or ϕ, ψ, χ by x, y, z does not vanish. Let $\phi'_x \neq 0$ be that derivative. Then $u = \phi(x, y, z)$ may be solved as $x = \omega(u, y, z)$ and the result may be substituted in v and w as

$$v = \psi(x, y, z) = \psi(\omega, y, z), \qquad w = \chi(x, y, z) = \chi(\omega, y, z).$$

Next the Jacobian of v and w relative to y and z may be written as

$$\begin{vmatrix} \dfrac{\partial v}{\partial y} & \dfrac{\partial w}{\partial y} \\ \dfrac{\partial v}{\partial z} & \dfrac{\partial w}{\partial z} \end{vmatrix} = \begin{vmatrix} \psi'_x\dfrac{\partial x}{\partial y} + \psi'_y & \chi'_x\dfrac{\partial x}{\partial y} + \chi'_y \\ \psi'_x\dfrac{\partial x}{\partial z} + \psi'_z & \chi'_x\dfrac{\partial x}{\partial z} + \chi'_z \end{vmatrix}$$

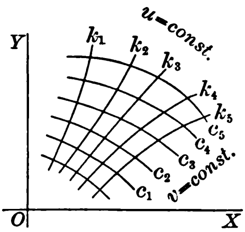
$$= \begin{vmatrix} \psi'_y & \chi'_y \\ \psi'_z & \chi'_z \end{vmatrix} + \psi'_x\begin{vmatrix} -\phi'_y/\phi'_x & \chi'_y \\ -\phi'_z/\phi'_x & \chi'_z \end{vmatrix} + \chi'_x\begin{vmatrix} \psi'_y & -\phi'_y/\phi'_x \\ \psi'_z & -\phi'_z/\phi'_x \end{vmatrix}$$

$$= \frac{1}{\phi'_x}\left[\phi'_x\begin{vmatrix} \psi'_y & \chi'_y \\ \psi'_z & \chi'_z \end{vmatrix} + \psi'_x\begin{vmatrix} \chi'_y & \phi'_y \\ \chi'_z & \phi'_z \end{vmatrix} + \chi'_x\begin{vmatrix} \phi'_y & \psi'_y \\ \phi'_z & \psi'_z \end{vmatrix}\right] = \frac{J}{\phi'_x}.$$

As J vanishes identically, the Jacobian of v and w expressed as functions of y, z, and u vanishes. Hence by the case previously discussed there is a functional relation $F(v, w) = 0$ independent of y, z; and as v, w now contain u, this relation may be considered as a functional relation between u, v, w.

63. If in (22) the variables u, v be assigned constant values, the equations define two curves, and if u, v be assigned a series of such values, the equations (22) define a network of curves in some part of the

xy-plane. If there is a functional relation $u = F(v)$, that is, if the Jacobian vanishes identically, a constant value of v implies a constant value of u and hence the locus for which v is constant is also a locus for which u is constant; the set of v-curves coincides with the set of u-curves and no true network is formed. This case is uninteresting. Let it be assumed that the Jacobian does not vanish identically and even that it does not vanish for any point (x, y) of a certain region of the xy-plane. The indications of § 60 are that the equations (22) may then be solved for x, y in terms of u, v at any point of the region and that there is a pair of the curves through each point. It is then proper to consider (u, v) as the coördinates of the points in the region. To any point there correspond not only the rectangular coördinates (x, y) but also the *curvilinear coördinates* (u, v).



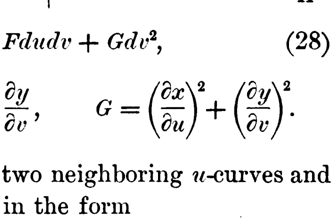
The equations connecting the rectangular and curvilinear coördinates may be taken in either of the two forms

$$u = \phi(x, y), \qquad v = \psi(x, y) \quad \text{or} \quad x = f(u, v), \qquad y = g(u, v), \qquad (22')$$

each of which are the solutions of the other. The Jacobians

$$J\left(\frac{u, v}{x, y}\right) \cdot J\left(\frac{x, y}{u, v}\right) = 1 \qquad (27)$$

are reciprocal each to each; and this relation may be regarded as the analogy of the relation (4) of § 2 for the case of the function $y = \phi(x)$ and the solution $x = f(y) = \phi^{-1}(y)$ in the case of a single variable. The *differential of arc* is



$$ds^2 = dx^2 + dy^2 = Edu^2 + 2\,Fdudv + Gdv^2, \qquad (28)$$

$$E = \left(\frac{\partial x}{\partial u}\right)^2 + \left(\frac{\partial y}{\partial u}\right)^2, \qquad F = \frac{\partial x}{\partial u}\frac{\partial x}{\partial v} + \frac{\partial y}{\partial u}\frac{\partial y}{\partial v}, \qquad G = \left(\frac{\partial x}{\partial u}\right)^2 + \left(\frac{\partial y}{\partial v}\right)^2.$$

The *differential of area* included between two neighboring u-curves and two neighboring v-curves may be written in the form

$$dA = J\left(\frac{x, y}{u, v}\right) dudv = dudv \div J\left(\frac{u, v}{x, y}\right). \qquad (29)$$

These statements will now be proved in detail.

To prove (27) write out the Jacobians at length and reduce the result.

$$J\left(\frac{u,\ v}{x,\ y}\right)J\left(\frac{x,\ y}{u,\ v}\right)=\begin{vmatrix}\frac{\partial u}{\partial x} & \frac{\partial v}{\partial x}\\ \frac{\partial u}{\partial y} & \frac{\partial v}{\partial y}\end{vmatrix}\cdot\begin{vmatrix}\frac{\partial x}{\partial u} & \frac{\partial y}{\partial u}\\ \frac{\partial x}{\partial v} & \frac{\partial y}{\partial v}\end{vmatrix}$$

$$=\begin{vmatrix}\frac{\partial u}{\partial x}\frac{\partial x}{\partial u}+\frac{\partial v}{\partial x}\frac{\partial x}{\partial v} & \frac{\partial u}{\partial x}\frac{\partial y}{\partial u}+\frac{\partial v}{\partial x}\frac{\partial y}{\partial v}\\ \frac{\partial u}{\partial y}\frac{\partial x}{\partial u}+\frac{\partial v}{\partial y}\frac{\partial x}{\partial v} & \frac{\partial u}{\partial y}\frac{\partial y}{\partial u}+\frac{\partial v}{\partial y}\frac{\partial y}{\partial v}\end{vmatrix}=\begin{vmatrix}1 & 0\\ 0 & 1\end{vmatrix}=1,$$

where the rule for multiplying determinants has been applied and the reduction has been made by (15), Ex. 9 above, and similar formulas. If the rule for multiplying determinants is unfamiliar, the Jacobians may be written and multiplied without that notation and the reduction may be made by the same formulas as before.

To establish the formula for the differential of arc it is only necessary to write the total differentials of dx and dy, to square and add, and then collect. To obtain the differential area between four adjacent curves consider the triangle determined by $(u,\ v)$, $(u + du,\ v)$, $(u,\ v + dv)$, which is half that area, and double the result. The determinantal form of the area of a triangle is the best to use.

$$dA = 2\cdot\frac{1}{2}\begin{vmatrix}d_ux & d_uy\\ d_vx & d_vy\end{vmatrix}=\begin{vmatrix}\frac{\partial x}{\partial u}du & \frac{\partial y}{\partial u}du\\ \frac{\partial x}{\partial v}dv & \frac{\partial y}{\partial v}\partial v\end{vmatrix}=\begin{vmatrix}\frac{\partial x}{\partial u} & \frac{\partial y}{\partial u}\\ \frac{\partial x}{\partial v} & \frac{\partial y}{\partial v}\end{vmatrix}dudv.$$

The subscripts on the differentials indicate which variable changes; thus d_ux, d_uy are the coördinates of $(u + du,\ v)$ relative to $(u,\ v)$. This method is easily extended to determine the analogous quantities in three dimensions or more. It may be noticed that the triangle does not look as if it were half the area (except for infinitesimals of higher order) in the figure; but see Ex. 12 below.

It should be remarked that as the differential of area dA is usually considered positive when du and dv are positive, it is usually better to replace J in (29) by its absolute value. Instead of regarding $(u,\ v)$ as curvilinear coördinates in the xy-plane, it is possible to plot them in their own uv-plane and thus to establish by (22′) a *transformation* of the xy-plane over onto the uv-plane. A small area in the xy-plane then becomes a small area in the uv-plane. If $J > 0$, the transformation is called direct; but if $J < 0$, the transformation is called perverted. The significance of the distinction can be made clear only when the question of the signs of areas has been treated. The transformation is called *conformal* when elements of arc in the neighborhood of a point in the xy-plane are proportional to the elements of arc in the neighborhood of the corresponding point in the uv-plane, that is, when

$$ds^2 = dx^2 + dy^2 = k\,(du^2 + dv^2) = kd\sigma^2. \tag{30}$$

For in this case any little triangle will be transformed into a little triangle similar to it, and hence angles will be unchanged by the transformation. That the transformation be conformal requires that $F = 0$ and $E = G$. It is not necessary that $E = G = k$ be constants; the ratio of similitude may be different for different points.

64. There remains outstanding the proof that equations may be solved in the neighborhood of a point at which the Jacobian does not vanish. The fact was indicated in § 60 and used in § 63.

THEOREM. Let p equations in $n + p$ variables be given, say,

$$F_1(x_1, x_2, \cdots, x_{n+p}) = 0, \qquad F_2 = 0, \cdots, F_p = 0. \tag{31}$$

Let the p functions be soluble for $x_{10}, x_{20}, \cdots, x_{p0}$ when a particular set $x_{(p+1)_0}, \cdots, x_{(n+p)_0}$ of the other n variables are given. Let the functions and their first derivatives be continuous in all the $n + p$ variables in the neighborhood of $(x_{10}, x_{20}, \cdots, x_{(n+p)_0})$. Let the Jacobian of the functions with respect to $x_1, x_2, \cdots, x_p$,

$$J\left(\frac{F_1, \cdots, F_p}{x_1, \cdots, x_p}\right) = \begin{vmatrix} \dfrac{\partial F_1}{\partial x_1} \cdots \dfrac{\partial F_p}{\partial x_1} \\ \vdots \qquad \vdots \\ \dfrac{\partial F_1}{\partial x_p} \cdots \dfrac{\partial F_p}{\partial x_p} \end{vmatrix}_{x_{10}, \cdots, x_{(n+p)_0}} \neq 0, \tag{32}$$

fail to vanish for the particular set mentioned. Then the p equations may be solved for the p variables $x_1, x_2, \cdots, x_p$, and the solutions will be continuous, unique, and differentiable with continuous first partial derivatives for all values of $x_{p+1}, \cdots, x_{n+p}$ sufficiently near to the values $x_{(p+1)_0}, \cdots, x_{(n+p)_0}$.

THEOREM. The necessary and sufficient condition that a functional relation exist between p functions of p variables is that the Jacobian of the functions with respect to the variables shall vanish identically, that is, for all values of the variables.

The proofs of these theorems will naturally be given by mathematical induction. Each of the theorems has been proved in the simplest cases and it remains only to show that the theorems are true for p functions in case they are for $p - 1$. Expand the determinant J.

$$J = J_1 \frac{\partial F_1}{\partial x_1} + J_2 \frac{\partial F_1}{\partial x_2} + \cdots + J_p \frac{\partial F_1}{\partial x_p}, \qquad J_1, \cdots, J_p, \text{ minors.}$$

For the first theorem $J \neq 0$ and hence at least one of the minors $J_1, \cdots, J_p$ must fail to vanish. Let that one be J_1, which is the Jacobian of $F_2, \cdots, F_p$ with respect to $x_2, \cdots, x_p$. By the assumption that the theorem holds for the case $p - 1$, these $p - 1$ equations may be solved for $x_2, \cdots, x_p$ in terms of the $n + 1$ variables x_1,

$x_{p+1}, \cdots, x_{n+p}$, and the results may be substituted in F_1. It remains to show that $F_1 = 0$ is soluble for x_1. Now

$$\frac{dF_1}{dx_1} = \frac{\partial F_1}{\partial x_1} + \frac{\partial F_1}{\partial x_2}\frac{\partial x_2}{\partial x_1} + \cdots + \frac{\partial F_1}{\partial x_p}\frac{\partial x_p}{\partial x_1} = J/J_1 \neq 0. \tag{32'}$$

For the derivatives of $x_2, \cdots, x_p$ with respect to x_1 are obtained from the equations

$$0 = \frac{\partial F_2}{\partial x_1} + \frac{\partial F_2}{\partial x_2}\frac{\partial x_2}{\partial x_1} + \cdots + \frac{\partial F_2}{\partial x_p}\frac{\partial x_p}{\partial x_1}, \quad \cdots, \quad 0 = \frac{\partial F_p}{\partial x_1} + \frac{\partial F_p}{\partial x_2}\frac{\partial x_2}{\partial x_1} + \cdots + \frac{\partial F_p}{\partial x_p}\frac{\partial x_p}{\partial x_1}$$

resulting from the differentiation of $F_2 = 0, \cdots, F_p = 0$ with respect to x_1. The derivative $\partial x_i/\partial x_1$ is therefore merely J_i/J_1, and hence $dF_1/dx_1 = J/J_1$ and does not vanish. The equation therefore may be solved for x_1 in terms of $x_{p+1}, \cdots, x_{n+p}$, and this result may be substituted in the solutions above found for $x_2, \cdots, x_p$. Hence the equations have been solved for $x_1, x_2, \cdots, x_p$ in terms of $x_{p+1}, \cdots, x_{n+p}$ and the theorem is proved.

For the second theorem the procedure is analogous to that previously followed. If there is a relation $F(u_1, \cdots, u_p) = 0$ between the p functions

$$u_1 = \phi_1(x_1, \cdots, x_p), \cdots, \qquad u_p = \phi_p(x_1, \cdots, x_p),$$

differentiation with respect to $x_1, \cdots, x_p$ gives p equations from which the derivatives of F by $u_1, \cdots, u_p$ may be eliminated and $J\left(\frac{u_1, \cdots, u_p}{x_1, \cdots, x_p}\right) = 0$ becomes the condition desired. If conversely this Jacobian vanishes identically and it be assumed that one of the derivatives of u_i by x_j, say $\partial u_1/\partial x_1$, does not vanish, then the solution $x_1 = \omega(u_1, x_2, \cdots, x_p)$ may be effected and the result may be substituted in $u_2, \cdots, u_p$. The Jacobian of $u_2, \cdots, u_p$ with respect to $x_2, \cdots, x_p$ will then turn out to be $J \div \partial u_1/\partial x_1$ and will vanish because J vanishes. Now, however, only $p - 1$ functions are involved, and hence if the theorem is true for $p - 1$ functions it must be true for p functions.

EXERCISES

1. If $u = ax + by + c$ and $v = a'x + b'y + c'$ are functionally dependent, the lines $u = 0$ and $v = 0$ are parallel; and conversely.

2. Prove $x + y + z$, $xy + yz + zx$, $x^2 + y^2 + z^2$ functionally dependent.

3. If $u = ax + by + cz + d$, $v = a'x + b'y + c'z + d'$, $w = a''x + b''y + c''z + d''$ are functionally dependent, the planes $u = 0$, $v = 0$, $w = 0$ are parallel to a line.

4. In what senses are $\frac{\partial v}{\partial y}$ and ψ_y' of (24′) and $\frac{dF_1}{dx_1}$ and $\frac{\partial F_1}{\partial x_1}$ of (32′) partial or total derivatives? Are not the two sets completely analogous?

5. Given (25), suppose $\begin{vmatrix} \psi_y' & \chi_y' \\ \psi_z' & \chi_z' \end{vmatrix} \neq 0$. Solve $v = \psi$ and $w = \chi$ for y and z, substitute in $u = \phi$, and prove $\partial u/\partial x = J \div \begin{vmatrix} \psi_y' & \chi_y' \\ \psi_z' & \chi_z' \end{vmatrix}$.

6. If $u = u(x, y)$, $v = v(x, y)$, and $x = x(\xi, \eta)$, $y = y(\xi, \eta)$, prove

$$J\left(\frac{u, v}{x, y}\right) J\left(\frac{x, y}{\xi, \eta}\right) = J\left(\frac{u, v}{\xi, \eta}\right). \tag{27'}$$

State the extension to any number of variables. How may (27′) be used to prove (27)? Again state the extension to any number of variables.

7. Prove $dV = J\left(\frac{x,\ y,\ z}{u,\ v,\ w}\right) dudvdw = dudvdw \div J\left(\frac{u,\ v,\ w}{x,\ y,\ z}\right)$ is the element of volume in space with curvilinear coördinates u, v, w = consts.

8. In what parts of the plane can $u = x^2 + y^2$, $v = xy$ not be used as curvilinear coördinates? Express ds^2 for these coördinates.

9. Prove that $2u = x^2 - y^2$, $v = xy$ is a conformal transformation.

10. Prove that $x = \frac{u}{u^2 + v^2}$, $y = \frac{v}{u^2 + v^2}$ is a conformal transformation.

11. Define conformal transformation in space. If the transformation

$$x = au + bv + cw, \qquad y = a'u + b'v + c'w, \qquad z = a''u + b''v + c''w$$

is conformal, is it orthogonal? See Ex. 10 (ζ), p. 100.

12. Show that the areas of the triangles whose vertices are

$$(u,\ v),\ (u + du,\ v),\ (u,\ v + dv) \quad \text{and} \quad (u + du,\ v + dv),\ (u + du,\ v),\ (u,\ v + dv)$$

are infinitesimals of the same order, as suggested in § 63.

13. Would the condition $F = 0$ in (28) mean that the set of curves u = const. were perpendicular to the set v = const.?

14. Express E, F, G in (28) in terms of the derivatives of u, v by x, y.

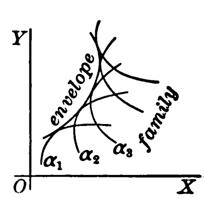
15. If $x = \phi(s,\ t)$, $y = \psi(s,\ t)$, $z = \chi(s,\ t)$ are the parametric equations of a surface (from which s, t could be eliminated to obtain the equation between x, y, z), show

$$\frac{\partial z}{\partial x} = J\left(\frac{\chi,\ \psi}{s,\ t}\right) \div J\left(\frac{\phi,\ \psi}{s,\ t}\right) \quad \text{and find} \quad \frac{\partial z}{\partial y}.$$

65. Envelopes of curves and surfaces. Let the equation $F(x, y, \alpha) = 0$ be considered as representing a family of curves where the different curves of the family are obtained by assigning different values to the parameter α. Such families are illustrated by

$$(x - \alpha)^2 + y^2 = 1 \quad \text{and} \quad \alpha x + y/\alpha = 1, \tag{33}$$

which are circles of unit radius centered on the x-axis and lines which cut off the area $\frac{1}{2}\alpha^2$ from the first quadrant. As α changes, the circles remain always tangent to the two lines $y = \pm 1$ and the point of tangency traces those lines. Again, as α changes, the lines (33) remain tangent to the hyperbola $xy = k$, owing to the property of the hyperbola that a tangent forms a triangle of constant area with the asymptotes. The lines $y = \pm 1$ are called the *envelope* of the system of circles and the hyperbola $xy = k$ the envelope of the set of lines. In general, *if there is a curve to which the curves of a family $F(x, y, \alpha) = 0$ are tangent and if the point of tangency describes that curve as α varies, the curve is called*



the envelope (or part of the envelope if there are several such curves) *of the family* $F(x, y, \alpha) = 0$. Thus any curve may be regarded as the envelope of its tangents or as the envelope of its circles of curvature.

To find the equations of the envelope note that by definition the enveloping curves of the family $F(x, y, \alpha) = 0$ are tangent to the envelope and that the point of tangency moves along the envelope as α varies. The equation of the envelope may therefore be written

$$x = \phi(\alpha), \qquad y = \psi(\alpha) \quad \text{with} \quad F(\phi, \psi, \alpha) = 0, \tag{34}$$

where the first equations express the dependence of the points on the envelope upon the parameter α and the last equation states that each point of the envelope lies also on some curve of the family $F(x, y, \alpha) = 0$. Differentiate (34) with respect to α. Then

$$F'_x\phi'(\alpha) + F'_y\psi'(\alpha) + F'_\alpha = 0. \tag{35}$$

Now if the point of contact of the envelope with the curve $F = 0$ is an ordinary point of that curve, the tangent to the curve is

$$F'_x(x - x_0) + F'_y(y - y_0) = 0; \quad \text{and} \quad F'_x\phi' + F'_y\psi' = 0,$$

since the tangent direction $dy : dx = \psi' : \phi'$ along the envelope is by definition identical with that along the enveloping curve; and if the point of contact is a singular point for the enveloping curve, $F'_x = F'_y = 0$. Hence in either case $F'_\alpha = 0$.

Thus *for points on the envelope the two equations*

$$F(x, y, \alpha) = 0, \qquad F'_\alpha(x, y, \alpha) = 0 \tag{36}$$

are satisfied and the equation of the envelope of the family $F = 0$ *may be found by solving* (36) *to find the parametric equations* $x = \phi(\alpha)$, $y = \psi(\alpha)$ *of the envelope or by eliminating* α *between* (36) *to find the equation of the envelope in the form* $\Phi(x, y) = 0$. It should be remarked that the locus found by this process may contain other curves than the envelope. For instance if the curves of the family $F = 0$ have singular points and if $x = \phi(\alpha)$, $y = \psi(\alpha)$ be the locus of the singular points as α varies, equations (34), (35) still hold and hence (36) also. The rule for finding the envelope therefore finds also the locus of singular points. Other extraneous factors may also be introduced in performing the elimination. It is therefore important to test graphically or analytically the solution obtained by applying the rule.

As a first example let the envelope of $(x - \alpha)^2 + y^2 = 1$ be found.

$$F(x, y, \alpha) = (x - \alpha)^2 + y^2 - 1 = 0, \qquad F'_\alpha = -2(x - \alpha) = 0.$$

The elimination of α from these equations gives $y^2 - 1 = 0$ and the solution for α gives $x = \alpha$, $y = \pm 1$. The loci indicated as envelopes are $y = \pm 1$. It is geometrically evident that these are really envelopes and not extraneous factors. But as a second example consider $\alpha x + y/\alpha = 1$. Here

$$F(x, y, \alpha) = \alpha x + y/\alpha - 1 = 0, \qquad F'_\alpha = x - y/\alpha^2 = 0.$$

The solution is $y = \alpha/2$, $x = 1/2\alpha$, which gives $xy = \frac{1}{4}$. This is the envelope; it could not be a locus of singular points of $F = 0$ as there are none. Suppose the elimination of α be made by Sylvester's method as

$$\begin{matrix} -y/\alpha^2 & +0/\alpha & +x+0\alpha = 0 \\ 0/\alpha^2 & -y/\alpha & +0+x\alpha = 0 \\ y/\alpha^2 & -1/\alpha & +x+0\alpha = 0 \\ 0/\alpha^2 & +y/\alpha & -1+x\alpha = 0 \end{matrix} \quad \text{and} \quad \begin{vmatrix} -y & 0 & x & 0 \\ 0 & -y & 0 & x \\ y & -1 & x & 0 \\ 0 & y & -1 & x \end{vmatrix} = 0;$$

the reduction of the determinant gives $xy(4xy - 1) = 0$ as the eliminant, and contains not only the envelope $4xy = 1$, but the factors $x = 0$ and $y = 0$ which are obviously extraneous.

As a third problem find the envelope of a line of which the length intercepted between the axes is constant. The necessary equations are

$$\frac{x}{\alpha} + \frac{y}{\beta} = 1, \qquad \alpha^2 + \beta^2 = K^2, \qquad \frac{x}{\alpha^2}d\alpha + \frac{y}{\beta^2}d\beta = 0, \qquad \alpha d\alpha + \beta d\beta = 0.$$

Two parameters α, β connected by a relation have been introduced; both equations have been differentiated totally with respect to the parameters; and the problem is to eliminate $\alpha, \beta, d\alpha, d\beta$ from the equations. In this case it is simpler to carry both parameters than to introduce the radicals which would be required if only one parameter were used. The elimination of $d\alpha, d\beta$ from the last two equations gives $x : y = \alpha^3 : \beta^3$ or $\sqrt[3]{x} : \sqrt[3]{y} = \alpha : \beta$. From this and the first equation,

$$\frac{1}{\alpha} = \frac{1}{x^{\frac{1}{3}}(x^{\frac{2}{3}} + y^{\frac{2}{3}})}, \quad \frac{1}{\beta} = \frac{1}{y^{\frac{1}{3}}(x^{\frac{2}{3}} + y^{\frac{2}{3}})}, \quad \text{and hence} \quad x^{\frac{2}{3}} + y^{\frac{2}{3}} = K^{\frac{2}{3}}.$$

66. Consider two neighboring curves of $F(x, y, \alpha) = 0$. Let (x_0, y_0) be an ordinary point of $\alpha = \alpha_0$ and $(x_0 + dx, y_0 + dy)$ of $\alpha_0 + d\alpha$. Then

$$\begin{aligned} F(x_0 + dx, y_0 + dy, \alpha_0 + d\alpha) &- F(x_0, y_0, \alpha_0) \\ &= F'_x dx + F'_y dy + F'_\alpha d\alpha = 0 \end{aligned} \tag{37}$$

holds except for infinitesimals of higher order. The distance from the point on $\alpha_0 + d\alpha$ to the tangent to α_0 at (x_0, y_0) is

$$\frac{F'_x dx + F'_y dy}{\pm\sqrt{F'^2_x + F'^2_y}} = \frac{\pm F'_\alpha d\alpha}{\sqrt{F'^2_x + F'^2_y}} = dn \tag{38}$$

except for infinitesimals of higher order. This distance is of the first order with $d\alpha$, and the normal derivative $d\alpha/dn$ of § 48 is finite except when $F'_\alpha = 0$. The distance is of higher order than $d\alpha$, and $d\alpha/dn$ is infinite or $dn/d\alpha$ is zero when $F'_\alpha = 0$. It appears therefore that *the envelope is the locus of points at which the distance between two neighboring curves is of higher order than $d\alpha$.* This is also apparent geometrically from the fact that the distance from a point on a curve to the tangent to the curve at a neighboring point is of higher order (§ 36). Singular points have been ruled out because (38) becomes indeterminate. In general the locus of singular points is not tangent to the curves of the family and is not an envelope but an extraneous factor; in exceptional cases this locus is an envelope.

If two neighboring curves $F(x, y, \alpha) = 0$, $F(x, y, \alpha + \Delta\alpha) = 0$ intersect, their point of intersection satisfies both of the equations, and hence also the equation

$$\frac{1}{\Delta\alpha}[F(x, y, \alpha + \Delta\alpha) - F(x, y, \alpha)] = F'_\alpha(x, y, \alpha + \theta\Delta\alpha) = 0.$$

If the limit be taken for $\Delta\alpha \doteq 0$, the limiting position of the intersection satisfies $F'_\alpha = 0$ and hence may lie on the envelope, and will lie on the envelope if the common point of intersection is remote from singular points of the curves $F(x, y, \alpha) = 0$. This idea of an *envelope as the limit of points in which neighboring curves of the family intersect* is valuable. It is sometimes taken as the definition of the envelope. But, unless imaginary points of intersection are considered, it is an inadequate definition; for otherwise $y = (x - \alpha)^3$ would have no envelope according to the definition (whereas $y = 0$ is obviously an envelope) and a curve could not be regarded as the envelope of its osculating circles.

Care must be used in applying the rule for finding an envelope. Otherwise not only may extraneous solutions be mistaken for the envelope, but the envelope may be missed entirely. Consider

$$y - \sin \alpha x = 0 \quad \text{or} \quad \alpha - x^{-1} \sin^{-1} y = 0, \tag{39}$$

where the second form is obtained by solution and contains a multiple valued function. These two families of curves are identical, and it is geometrically clear that they have an envelope, namely $y = \pm 1$. This is precisely what would be found on applying the rule to the first of (39); but if the rule be applied to the second of (39), it is seen that $F'_\alpha = 1$, which does not vanish and hence indicates no envelope. The whole matter should be examined carefully in the light of implicit functions.

Hence let $F(x, y, \alpha) = 0$ be a continuous single valued function of the three variables (x, y, α) and let its derivatives F'_x, F'_y, F'_α exist and be continuous. Consider the behavior of the curves of the family near a point (x_0, y_0) of the curve for $\alpha = \alpha_0$ provided that (x_0, y_0) is an ordinary (nonsingular) point of the curve and that the derivative $F'_\alpha(x_0, y_0, \alpha_0)$ does not vanish. As $F'_\alpha \neq 0$ and either $F'_x \neq 0$ or $F'_y \neq 0$ for (x_0, y_0, α_0), it is possible to surround (x_0, y_0) with a region so small that $F(x, y, \alpha) = 0$ may be solved for $\alpha = f(x, y)$ which will be single valued and differentiable; and the region may further be taken so small that F'_x or F'_y remains different from 0 throughout the region. Then through every point of the region there is one and only one curve $\alpha = f(x, y)$ and the curves have no singular points within the region. In particular no two curves of the family can be tangent to each other within the region.

Furthermore, in such a region there is no envelope. For let any curve which traverses the region be $x = \phi(t)$, $y = \psi(t)$. Then

$$\alpha(t) = f(\phi(t), \psi(t)), \qquad \alpha'(t) = f'_x\phi'(t) + f'_y\psi'(t).$$

Along any curve $\alpha = f(x, y)$ the equation $f'_x dx + f'_y dy = 0$ holds, and if $x = \phi(t)$, $y = \psi(t)$ be tangent to this curve, $dy = dx = \psi' : \phi'$ and $\alpha'(t) = 0$ or $\alpha = \text{const.}$ Hence the only curve which has at each point the direction of the curve of the family through that point is a curve which coincides throughout with some curve of the family and is tangent to no other member of the family. Hence there is no envelope. The result is that an envelope can be present only when $F'_\alpha = 0$ or when $F'_x = F'_y = 0$, and this latter case has been seen to be included in the condition $F'_\alpha = 0$. If $F(x, y, \alpha)$ were not single valued but the branches were separable, the same conclusion would hold. Hence in case $F(x, y, \alpha)$ is not single valued the loci over which two or more values become inseparable must be added to those over which $F'_\alpha = 0$ in order to insure that all the loci which may be envelopes are taken into account.

67. The preceding considerations apply with so little change to other cases of envelopes that the facts will merely be stated without proof. Consider a family of surfaces $F(x, y, z, \alpha, \beta) = 0$ depending on two parameters. The envelope may be defined by the property of tangency as in § 65; and *the conditions for an envelope would be*

$$F(x, y, z, \alpha, \beta) = 0, \qquad F'_\alpha = 0, \qquad F'_\beta = 0. \tag{40}$$

These three equations may be solved to express the envelope as

$$x = \phi(\alpha, \beta), \qquad y = \psi(\alpha, \beta), \qquad z = \chi(\alpha, \beta)$$

parametrically in terms of α, β; or the two parameters may be eliminated and the envelope may be found as $\Phi(x, y, z) = 0$. In any case extraneous loci may be introduced and the results of the work should therefore be tested, which generally may be done at sight.

It is also possible to determine the distance from the tangent plane of one surface to the neighboring surfaces as

$$\frac{F'_x dx + F'_y dy + F'_z dz}{\sqrt{F'^2_x + F'^2_y + F'^2_z}} = \frac{F'_\alpha d\alpha + F'_\beta d\beta}{\sqrt{F'^2_x + F'^2_y + F'^2_z}} = dn, \tag{41}$$

and to define the envelope as the locus of points such that this distance is of higher order than $|d\alpha| + |d\beta|$. The equations (40) would then also follow. This definition would apply only to ordinary points of the surfaces of the family, that is, to points for which not all the derivatives F'_x, F'_y, F'_z vanish. But as the elimination of α, β from (40) would give an equation which included the loci of these singular points, there would be no danger of losing such loci in the rare instances where they, too, happened to be tangent to the surfaces of the family.

The application of implicit functions as in § 66 could also be made in this case and would show that no envelope could exist in regions where no singular points occurred and where either F'_α or F'_β failed to vanish. This work could be based either on the first definition involving tangency directly or on the second definition which involves tangency indirectly in the statements concerning infinitesimals of higher order. It may be added that if $F(x, y, z, \alpha, \beta) = 0$ were not single valued, the surfaces over which two values of the function become inseparable should be added as possible envelopes.

A family of surfaces $F(x, y, z, \alpha) = 0$ depending on a single parameter may have an envelope, and *the envelope is found from*

$$F(x, y, z, \alpha) = 0, \qquad F'_\alpha(x, y, z, \alpha) = 0 \tag{42}$$

by the elimination of the single parameter. The details of the deduction of the rule will be omitted. If two neighboring surfaces intersect, the limiting position of the curve of intersection lies on the envelope and the envelope is the surface generated by this curve as α varies. The surfaces of the family touch the envelope not at a point merely but along these curves. The curves are called *characteristics* of the family. In the case where consecutive surfaces of the family do not intersect in a real curve it is necessary to fall back on the conception of imaginaries or on the definition of an envelope in terms of tangency or infinitesimals; the characteristic curves are still the curves along which the surfaces of the family are in contact with the envelope and along which two consecutive surfaces of the family are distant from each other by an infinitesimal of higher order than $d\alpha$.

A particular case of importance is the envelope of a plane which depends on one parameter. The equations (42) are then

$$Ax + By + Cz + D = 0, \qquad A'x + B'y + C'z + D' = 0, \tag{43}$$

where A, B, C, D are functions of the parameter and differentiation with respect to it is denoted by accents. The case where the plane moves parallel to itself or turns about a line may be excluded as trivial. As the intersection of two planes is a line, the characteristics of the system are straight lines, the envelope is a *ruled surface*, and *a plane tangent to the surface at one point of the lines is tangent to the surface throughout the whole extent of the line.* Cones and cylinders are examples of this sort of surface. Another example is the surface enveloped by the osculating planes of a curve in space; for the osculating plane depends on only one parameter. As the osculating plane (§ 41) may be regarded as passing through three consecutive points of the curve, two consecutive osculating planes may be considered as having two consecutive points of the curve in common and hence the characteristics are the tangent lines to the curve. Surfaces which are the envelopes of a plane which depends on a single parameter are called *developable surfaces*.

A family of curves dependent on two parameters as

$$F(x, y, z, \alpha, \beta) = 0, \qquad G(x, y, z, \alpha, \beta) = 0 \tag{44}$$

is called a *congruence of curves*. The curves may have an envelope, that is, there may be a surface to which the curves are tangent and which may be regarded as the locus of their points of tangency. The envelope is obtained by eliminating α, β from the equations

$$F = 0, \qquad G = 0, \qquad F'_\alpha G'_\beta - F'_\beta G'_\alpha = 0. \tag{45}$$

To see this, suppose that the third condition is not fulfilled. The equations (44) may then be solved as $\alpha = f(x, y, z)$, $\beta = g(x, y, z)$. Reasoning like that of § 66 now shows that there cannot possibly be an envelope in the region for which the solution is valid. It may therefore be inferred that the only possibilities for an envelope are contained in the equations (45). As various extraneous loci might be introduced in the elimination of α, β from (45) and as the solutions should therefore be tested individually, it is hardly necessary to examine the general question further. The envelope of a congruence of curves is called the *focal surface* of the congruence and the points of contact of the curves with the envelope are called the *focal points* on the curves.

EXERCISES

1. Find the envelopes of these families of curves. In each case test the answer or its individual factors and check the results by a sketch:

$$(\alpha)\ y = 2\alpha x + \alpha^4, \qquad (\beta)\ y^2 = \alpha(x - \alpha), \qquad (\gamma)\ y = \alpha x + k/\alpha,$$
$$(\delta)\ \alpha(y + \alpha)^2 = x^3, \qquad (\epsilon)\ y = \alpha(x + \alpha)^2, \qquad (\zeta)\ y^2 = \alpha(x - \alpha)^3.$$

2. Find the envelope of the ellipses $x^2/a^2 + y^2/b^2 = 1$ under the condition that (α) the sum of the axes is constant or (β) the area is constant.

3. Find the envelope of the circles whose center is on a given parabola and which pass through the vertex of the parabola.

4. Circles pass through the origin and have their centers on $x^2 - y^2 = c^2$. Find their envelope. *Ans.* A lemniscate.

5. Find the envelopes in these cases:

$$(\alpha)\ x + xy\alpha = \sin^{-1}xy, \qquad (\beta)\ x + \alpha = \operatorname{vers}^{-1}y + \sqrt{2y - y^2},$$
$$(\gamma)\ y + \alpha = \sqrt{1 - 1/x}.$$

6. Find the envelopes in these cases:

$$(\alpha)\ \alpha x + \beta y + \alpha\beta z = 1, \qquad (\beta)\ \frac{x}{\alpha} + \frac{y}{\beta} + \frac{z}{1 - \alpha - \beta} = 1,$$
$$(\gamma)\ \frac{x^2}{\alpha^2} + \frac{y^2}{\beta^2} + \frac{z^2}{\gamma^2} = 1 \text{ with } \alpha\beta\gamma = k^3.$$

7. Find the envelopes in Ex. 6 (α), (β) if $\alpha = \beta$ or if $\alpha = -\beta$.

8. Prove that the envelope of $F(x, y, z, \alpha) = 0$ is tangent to the surface along the whole characteristic by showing that the normal to $F(x, y, z, \alpha) = 0$ and to the eliminant of $F = 0$, $F'_\alpha = 0$ are the same, namely

$$F'_x : F'_y : F'_z \quad \text{and} \quad F'_x + F'_\alpha \frac{\partial \alpha}{\partial x} : F'_y + F'_\alpha \frac{\partial \alpha}{\partial y} : F'_z + F'_\alpha \frac{\partial \alpha}{\partial z},$$

where $\alpha(x, y, z)$ is the function obtained by solving $F'_\alpha = 0$. Consider the problem also from the point of view of infinitesimals and the normal derivative.

9. If there is a curve $x = \phi(\alpha)$, $y = \psi(\alpha)$, $z = \chi(\alpha)$ tangent to the curves of the family defined by $F(x, y, z, \alpha) = 0$, $G(x, y, z, \alpha) = 0$ in space, then that curve is called the envelope of the family. Show, by the same reasoning as in § 65 for the case of the plane, that the four conditions $F = 0$, $G = 0$, $F'_\alpha = 0$, $G'_\alpha = 0$ must be satisfied for an envelope; and hence infer that ordinarily a family of curves in space dependent on a single parameter has no envelope.

10. Show that the family $F(x, y, z, \alpha) = 0$, $F'_\alpha(x, y, z, \alpha) = 0$ of curves which are the characteristics of a family of surfaces has in general an envelope given by the three equations $F = 0$, $F'_\alpha = 0$, $F''_{\alpha\alpha} = 0$.

11. Derive the condition (45) for the envelope of a two-parametered family of curves from the idea of tangency, as in the case of one parameter.

12. Find the envelope of the normals to a plane curve $y = f(x)$ and show that the envelope is the locus of the center of curvature.

13. The locus of Ex. 12 is called the *evolute* of the curve $y = f(x)$. In these cases find the evolute as an envelope:

(α) $y = x^2$, (β) $x = a \sin t$, $y = b \cos t$, (γ) $2xy = a^2$,
(δ) $y^2 = 2mx$, (ϵ) $x = a(\theta - \sin\theta)$, $y = a(1 - \cos\theta)$, (ζ) $y = \cosh x$.

14. Given a surface $z = f(x, y)$. Construct the family of normal lines and find their envelope.

15. If rays of light issuing from a point in a plane are reflected from a curve in the plane, the angle of reflection being equal to the angle of incidence, the envelope of the reflected rays is called the *caustic* of the curve with respect to the point. Show that the caustic of a circle with respect to a point on its circumference is a cardioid.

16. The curve which is the envelope of the characteristic lines, that is, of the rulings, on the developable surface (43) is called the *cuspidal edge* of the surface. Show that the equations of this curve may be found parametrically in terms of the parameter of (43) by solving simultaneously

$$Ax + By + Cz + D = 0,\ A'x + B'y + C'z + D' = 0,\ A''x + B''y + C''z + D'' = 0$$

for x, y, z. Consider the exceptional cases of cones and cylinders.

17. The term "developable" signifies that a *developable surface may be developed or mapped on a plane in such a way that lengths of arcs on the surface become equal lengths in the plane*, that is, the map may be made without distortion of size or shape. In the case of cones or cylinders this map may be made by slitting the cone or cylinder along an element and rolling it out upon a plane. What is the analytic statement in this case? In the case of any developable surface with a cuspidal edge, the developable surface being the locus of all tangents to the cuspidal edge, the length of arc upon the surface may be written as $d\sigma^2 = (dt + ds)^2 + t^2ds^2/R^2$, where s denotes arc measured along the cuspidal edge and t denotes distance along the tangent line. This form of $d\sigma^2$ may be obtained geometrically by infinitesimal analysis or analytically from the equations

$$x = f(s) + tf'(s), \quad y = g(s) + tg'(s), \quad z = h(s) + th'(s)$$

of the developable surface of which $x = f(s)$, $y = g(s)$, $z = h(s)$ is the cuspidal edge. It is thus seen that $d\sigma^2$ is the same at corresponding points of all developable surfaces for which the radius of curvature R of the cuspidal edge is the same function of s without regard to the torsion; in particular the torsion may be zero and the developable may reduce to a plane.

18. Let the line $x = az + b$, $y = cz + d$ depend on one parameter so as to generate a ruled surface. By identifying this form of the line with (43) obtain by substitution the conditions

$$\begin{matrix} Aa + Bc + C = 0, & A'a + B'c + C' = 0 \\ Ab + Bd + D = 0, & A'b + B'd + D' = 0 \end{matrix} \quad \text{or} \quad \begin{matrix} Aa' + Bc' = 0 \\ Ab' + Bd' = 0 \end{matrix} \quad \text{or} \quad \begin{vmatrix} a' & c' \\ b' & d' \end{vmatrix} = 0$$

as the condition that the line generates a developable surface.

68. More differential geometry.

The representation

$$F(x, y, z) = 0, \quad \text{or} \quad z = f(x, y) \tag{46}$$

or

$$x = \phi(u, v), \qquad y = \psi(u, v), \qquad z = \chi(u, v)$$

of a surface may be taken in the unsolved, the solved, or the parametric form. The parametric form is equivalent to the solved form provided u, v be taken as x, y. The notation

$$p = \frac{\partial z}{\partial x}, \quad q = \frac{\partial z}{\partial y}, \quad r = \frac{\partial^2 z}{\partial x^2}, \quad s = \frac{\partial^2 z}{\partial x \partial y}, \quad t = \frac{\partial^2 z}{\partial y^2}$$

is adopted for the derivatives of z with respect to x and y. The application of Taylor's Formula to the solved form gives

$$\Delta z = ph + qk + \tfrac{1}{2}(rh^2 + 2\,shk + tk^2) + \cdots \tag{47}$$

with $h = \Delta x$, $k = \Delta y$. The linear terms $ph + qk$ constitute the differential dz and represent that part of the increment of z which would be obtained by replacing the surface by its tangent plane. Apart from infinitesimals of the third order, the distance from the tangent plane up or down to the surface along a parallel to the z-axis is given by the quadratic terms $\frac{1}{2}(rh^2 + 2\,shk + tk^2)$.

Hence if the quadratic terms at any point are a positive definite form (§ 55), the surface lies above its tangent plane and is concave up; but if the form is negative definite, the surface lies below its tangent plane and is concave down or convex up. If the form is indefinite but not singular, the surface lies partly above and partly below its tangent plane and may be called concavo-convex, that is, it is saddle-shaped. If the form is singular nothing can be definitely stated. These statements are merely generalizations of those of § 55 made for the case where the tangent plane is parallel to the xy-plane. It will be assumed in the further work of these articles that at least one of the derivatives r, s, t is not 0.

To examine more closely the behavior of a surface in the vicinity of a particular point upon it, let the xy-plane be taken in coincidence with the tangent plane at the point and let the point be taken as origin. Then Maclaurin's Formula is available.

$$\begin{aligned} z &= \tfrac{1}{2}(rx^2 + 2\,sxy + ty^2) + \text{terms of higher order} \\ &= \tfrac{1}{2}\rho^2(r\cos^2\theta + 2\,s\sin\theta\cos\theta + t\sin^2\theta) + \text{higher terms,} \end{aligned} \tag{48}$$

where (ρ, θ) are polar coördinates in the xy-plane. Then

$$\frac{1}{R} = r\cos^2\theta + 2\,s\sin\theta\cos\theta + t\sin^2\theta = \frac{d^2z}{d\rho^2} \div \left[1 + \left(\frac{dz}{d\rho}\right)^2\right]^{\frac{3}{2}} \tag{49}$$

is the curvature of a normal section of the surface. The sum of the curvatures in two normal sections which are in perpendicular planes may be obtained by giving θ the values θ and $\theta + \frac{1}{2}\pi$. This sum reduces to $r + t$ and is therefore independent of θ.

As the sum of the curvatures in two perpendicular normal planes is constant, the maximum and minimum values of the curvature will be found in perpendicular planes. These values of the curvature are called the *principal values* and their reciprocals are the *principal radii of curvature* and the sections in which they lie are the *principal sections.* If $s = 0$, the principal sections are $\theta = 0$ and $\theta = \frac{1}{2}\pi$; and conversely if the axes of x and y had been chosen in the tangent plane so as to be tangent to the principal sections, the derivative s would have vanished. The equation of the surface would then have taken the simple form

$$z = \tfrac{1}{2}(rx^2 + ty^2) + \text{higher terms.} \tag{50}$$

The principal curvatures would be merely r and t, and the curvature in any normal section would have had the form

$$\frac{1}{R} = \frac{\cos^2\theta}{R_1} + \frac{\sin^2\theta}{R_2} = r\cos^2\theta + t\sin^2\theta.$$

If the two principal curvatures have opposite signs, that is, if the signs of r and t in (50) are opposite, the surface is saddle-shaped. There are then two directions for which the curvature of a normal section vanishes, namely the directions of the lines

$$\theta = \pm\tan^{-1}\sqrt{-R_2/R_1} \quad \text{or} \quad \sqrt{|r|}\,x = \pm\sqrt{|t|}\,y.$$

These are called the *asymptotic directions.* Along these directions the surface departs from its tangent plane by infinitesimals of the third order, or higher order. If a curve is drawn on a surface so that at each point of the curve the tangent to the curve is along one of the asymptotic directions, the curve is called an *asymptotic curve or line* of the surface. As the surface departs from its tangent plane by infinitesimals of higher order than the second along an asymptotic line, the tangent plane to a surface at any point of an asymptotic line must be the osculating plane of the asymptotic line.

The character of a point upon a surface is indicated by the *Dupin indicatrix* of the point. The indicatrix is the conic

$$\frac{x^2}{R_1} + \frac{y^2}{R_2} = 1, \qquad \text{cf. } z = \tfrac{1}{2}(rx^2 + ty^2), \tag{51}$$

which has the principal directions as the directions of its axes and the square roots of the absolute values of the principal radii of curvature as the magnitudes of its axes. The conic may be regarded as similar to the conic in which a plane infinitely near the tangent plane cuts the surface when infinitesimals of order higher than the second are neglected. In case the surface is concavo-convex the indicatrix is a hyperbola and should be considered as either or both of the two conjugate hyperbolas that would arise from giving z positive or negative values in (51). The point on the surface is called elliptic, hyperbolic, or parabolic according as the indicatrix is an ellipse, a hyperbola, or a pair of lines, as happens when one of the principal curvatures vanishes. These classes of points correspond to the distinctions definite, indefinite, and singular applied to the quadratic form $rh^2 + 2shk + tk^2$.

Two further results are noteworthy. Any curve drawn on the surface differs from the section of its osculating plane with the surface by infinitesimals of higher order than the second. For as the osculating plane passes through three consecutive points of the curve, its intersection with the surface passes through the same three consecutive points and the two curves have contact of the second order. It follows that the radius of curvature of any curve on the surface is identical with that of the curve in which its osculating plane cuts the surface. The other result is *Meusnier's Theorem:* The radius of curvature of an oblique section of the surface at any point is the projection upon the plane of that section of the radius of curvature of the normal section which passes through the same tangent line. In other words, if the radius of curvature of a normal section is known, that of the oblique sections through the same tangent line may be obtained by multiplying it by the cosine of the angle between the plane normal to the surface and the plane of the oblique section.

The proof of Meusnier's Theorem may be given by reference to (48). Let the x-axis in the tangent plane be taken along the intersection with the oblique plane. Neglect infinitesimals of higher order than the second. Then

$$y = \phi(x) = \tfrac{1}{2}ax^2, \qquad z = \tfrac{1}{2}(rx^2 + 2sxy + ty^2) = \tfrac{1}{2}rx^2 \tag{48'}$$

will be the equations of the curve. The plane of the section is $az - ry = 0$, as may be seen by inspection. The radius of curvature of the curve in this plane may be found at once. For if u denote distance in the plane and perpendicular to the x-axis and if ν be the angle between the normal plane and the oblique plane $az - ry = 0$,

$$u = z \sec \nu = y \csc \nu = \tfrac{1}{2} r \sec \nu \cdot x^2 = \tfrac{1}{2} a \csc \nu \cdot x^2.$$

The form $u = \frac{1}{2} r \sec \nu \cdot x^2$ gives the curvature as $r \sec \nu$. But the curvature in the normal section is r by (48'). As the curvature in the oblique section is $\sec \nu$ times that in the normal section, the radius of curvature in the oblique section is $\cos \nu$ times that of the normal section. Meusnier's Theorem is thus proved.

69. These investigations with a special choice of axes give geometric properties of the surface, but do not express those properties in a convenient analytic form; for if a surface $z = f(x, y)$ is given, the transformation to the special axes is difficult. The idea of the indicatrix or its similar conic as the section of the surface by a plane near the tangent plane and parallel to it will, however, determine the general conditions readily. If in the expansion

$$\Delta z - dz = \tfrac{1}{2}(rh^2 + 2shk + tk^2) = \text{const.} \tag{52}$$

the quadratic terms be set equal to a constant, the conic obtained is the projection of the indicatrix on the xy-plane, or if (52) be regarded as a cylinder upon the xy-plane, the indicatrix (or similar conic) is the intersection of the cylinder with the tangent plane. As the character of the conic is unchanged by the projection, *the point on the surface is elliptic if $s^2 < rt$, hyperbolic if $s^2 > rt$, and parabolic if $s^2 = rt$.* Moreover if the indicatrix is hyperbolic, its asymptotes must project into the asymptotes of the conic (52), and hence if dx and dy replace h and k, the equation

$$rdx^2 + 2sdxdy + tdy^2 = 0 \tag{53}$$

may be regarded as *the differential equation of the projection of the asymptotic lines on the xy-plane.* If r, s, t be expressed as functions $f''_{xx}, f''_{xy}, f''_{yy}$ of (x, y) and (53) be factored, the integration of the two equations $M(x, y)\,dx + N(x, y)\,dy$ thus found will give the finite equations of the projections of the asymptotic lines and, taken with the equation $z = f(x, y)$, will give the curves on the surface.

To find the lines of curvature is not quite so simple; for it is necessary to determine the directions which are the projections of the axes of the indicatrix, and these are not the axes of the projected conic. Any radius of the indicatrix may be regarded as the intersection of the tangent plane and a plane perpendicular to the xy-plane through the radius of the projected conic. Hence

$$z - z_0 = p(x - x_0) + q(y - y_0), \qquad (x - x_0)k = (y - y_0)h$$

are the two planes which intersect in the radius that projects along the direction determined by h, k. The direction cosines

$$\frac{h : k : ph + qk}{\sqrt{h^2 + k^2 + (ph + qk)^2}} \quad \text{and} \quad h : k : 0 \tag{54}$$

are therefore those of the radius in the indicatrix and of its projection and they determine the cosine of the angle ϕ between the radius and its projection. The square of the radius in (52) is

$$h^2 + k^2, \quad \text{and} \quad (h^2 + k^2)\sec^2\phi = h^2 + k^2 + (ph + qk)^2$$

is therefore the square of the corresponding radius in the indicatrix. To determine the axes of the indicatrix, this radius is to be made a maximum or minimum subject to (52). With a multiplier λ,

$$h + ph + qk + \lambda(rh + sk) = 0, \qquad k + ph + qk + \lambda(sh + tk) = 0$$

are the conditions required, and the elimination of λ gives

$$h^2[s(1 + p^2) - pqr] + hk[t(1 + p^2) - r(1 + q^2)] - k^2[t(1 + q^2) - pqt] \doteq 0$$

as the equation that determines the projection of the axes. Or

$$\frac{(1 + p^2)\,dx + pqdy}{rdx + sdy} = \frac{pqdx + (1 + q^2)\,dy}{sdx + tdy} \tag{55}$$

is the differential equation of the projected lines of curvature.

In addition to the asymptotic lines and lines of curvature the *geodesic or shortest lines* on the surface are important. These, however, are better left for the methods of the calculus of variations (§ 159). The attention may therefore be turned to finding the value of the radius of curvature in any normal section of the surface.

A reference to (48) and (49) shows that the curvature is

$$\frac{1}{R} = \frac{2z}{\rho^2} = \frac{rh^2 + 2shk + tk^2}{\rho^2} = \frac{rh^2 + 2shk + tk^2}{h^2 + k^2}$$

in the special case. But in the general case the normal distance to the surface is $(\Delta z - dz)\cos\gamma$, with $\sec\gamma = \sqrt{1 + p^2 + q^2}$, instead of the $2z$ of the special case, and the radius ρ^2 of the special case becomes $\rho^2\sec^2\phi = h^2 + k^2 + (ph + qk)^2$ in the tangent plane. Hence

$$\frac{1}{R} = \frac{2(\Delta z - dz)\cos\gamma}{h^2 + k^2 + (ph + qk)^2} = \frac{rl^2 + 2slm + tm^2}{\sqrt{1 + p^2 + q^2}}, \tag{56}$$

where the direction cosines l, m of a radius in the tangent plane have been introduced from (54), is the general expression for the curvature of a normal section. The form

$$\frac{1}{R} = \frac{rh^2 + 2shk + tk^2}{h^2 + k^2 + (ph + qk)^2}\,\frac{1}{\sqrt{1 + p^2 + q^2}}, \tag{56'}$$

where the direction h, k of the projected radius remains, is frequently more convenient than (56) which contains the direction cosines l, m of the original direction in the tangent plane. Meusnier's Theorem may now be written in the form

$$\frac{\cos\nu}{R} = \frac{rl^2 + 2slm + tm^2}{\sqrt{1 + p^2 + q^2}}, \tag{57}$$

where ν is the angle between an oblique section and the tangent plane and where l, m are the direction cosines of the intersection of the planes.

The work here given has depended for its relative simplicity of statement upon the assumption of the surface (46) in solved form. It is merely a problem in implicit partial differentiation to pass from p, q, r, s, t to their equivalents in terms of F'_x, F'_y, F'_z or the derivatives of ϕ, ψ, χ by α, β.

EXERCISES

1. In (49) show $\frac{1}{R} = \frac{r+t}{2} + \frac{r-t}{2}\cos 2\theta + s\sin 2\theta$ and find the directions of maximum and minimum R. If R_1 and R_2 are the maximum and minimum values of R, show

$$\frac{1}{R_1} + \frac{1}{R_2} = r + t \quad \text{and} \quad \frac{1}{R_1}\frac{1}{R_2} = rt - s^2.$$

Half of the sum of the curvatures is called the *mean curvature;* the product of the curvatures is called the *total curvature.*

2. Find the mean curvature, the total curvature, and therefrom (by constructing and solving a quadratic equation) the principal radii of curvature at the origin:

$$(\alpha)\ z = xy, \qquad (\beta)\ z = x^2 + xy + y^2, \qquad (\gamma)\ z = x(x + y).$$

3. In the surfaces (α) $z = xy$ and (β) $z = 2x^2 + y^2$ find at (0, 0) the radius of curvature in the sections made by the planes

$$(\alpha)\ x + y = 0, \qquad (\beta)\ x + y + z = 0, \qquad (\gamma)\ x + y + 2z = 0,$$
$$(\delta)\ x - 2y = 0, \qquad (\epsilon)\ x - 2y + z = 0, \qquad (\zeta)\ x + 2y + \tfrac{1}{2}z = 0.$$

The oblique sections are to be treated by applying Meusnier's Theorem.

4. Find the asymptotic directions at (0, 0) in Exs. 2 and 3.

5. Show that *a developable surface is everywhere parabolic*, that is, that $rt - s^2 = 0$ at every point; and conversely. To do this consider the surface as the envelope of its tangent plane $z - p_0x - q_0y = z_0 - p_0x_0 - q_0y_0$, where p_0, q_0, x_0, y_0, z_0 are functions of a single parameter α. Hence show

$$J\left(\frac{p_0,\ q_0}{x_0,\ y_0}\right) = 0 = (rt - s^2)_0 \quad \text{and} \quad J\left(\frac{p_0,\ z_0 - p_0x_0 - q_0y_0}{x_0,\ y_0}\right) = y_0(s^2 - rt)_0.$$

The first result proves the statement; the second, its converse.

6. Find the differential equations of the asymptotic lines and lines of curvature on these surfaces:

$$(\alpha)\ z = xy, \qquad (\beta)\ z = \tan^{-1}(y/x), \qquad (\gamma)\ z^2 + y^2 = \cosh x, \qquad (\delta)\ xyz = 1.$$

7. Show that the mean curvature and total curvature are

$$\frac{1}{2}\left(\frac{1}{R_1} + \frac{1}{R_2}\right) = \frac{r(1 + q^2) + t(1 + p^2) - 2pqs}{2(1 + p^2 + q^2)^{\frac{3}{2}}}, \qquad \frac{1}{R_1R_2} = \frac{rt - s^2}{(1 + p^2 + q^2)^2}.$$

8. Find the principal radii of curvature at (1, 1) in Ex. 6.

9. An umbilic is a point of a surface at which the principal radii of curvature (and hence all radii of curvature for normal sections) are equal. Show that the conditions are $\frac{r}{1+p^2} = \frac{s}{pq} = \frac{t}{1+q^2}$ for an umbilic, and determine the umbilics of the ellipsoid with semiaxes a, b, c.

CHAPTER VI

COMPLEX NUMBERS AND VECTORS

70. Operators and operations. If an entity u is changed into an entity v by some law, the change may be regarded as an *operation* performed upon u, the *operand*, to convert it into v; and if f be introduced as the symbol of the operation, the result may be written as $v = fu$. For brevity the symbol f is often called an *operator*. Various sorts of operand, operator, and result are familiar. Thus if u is a positive number n, the application of the operator $\sqrt{\ }$ gives the square root; if u represents a range of values of a variable x, the expression $f(x)$ or fx denotes a function of x; if u be a function of x, the operation of differentiation may be symbolized by D and the result Du is the derivative; the symbol of definite integration $\int_a^b (*)d*$ converts a function $u(x)$ into a number; and so on in great variety.

The reason for making a short study of operators is that a considerable number of the concepts and rules of arithmetic and algebra may be so defined for operators themselves as to lead to a *calculus of operations* which is of frequent use in mathematics; the single application to the integration of certain differential equations (§ 95) is in itself highly valuable. The fundamental concept is that of a *product: If u is operated upon by f to give $fu = v$ and if v is operated upon by g to give $gv = w$, so that*

$$fu = v, \qquad gv = gfu = w, \qquad gfu = w, \tag{1}$$

then the operation indicated as gf which converts u directly into w is called the product of f by g. If the functional symbols sin and log be regarded as operators, the symbol log sin could be regarded as the product. The transformations of turning the xy-plane over on the x-axis, so that $x' = x$, $y' = -y$, and over the y-axis, so that $x' = -x$, $y' = y$, may be regarded as operations; the combination of these operations gives the transformation $x' = -x$, $y' = -y$, which is equivalent to rotating the plane through 180° about the origin.

The products of arithmetic and algebra satisfy the *commutative law* $gf = fg$, that is, the products of g by f and of f by g are equal. This is not true of operators in general, as may be seen from the fact that log sin x and sin log x are different. Whenever the order of the factors is immaterial, as in the case of the transformations just considered, the operators are said to be *commutative*. Another law of arithmetic and algebra is that when there are three or more factors in a product, the factors may be grouped at pleasure without altering the result, that is,

$$h(gf) = (hg)f = hgf. \tag{2}$$

This is known as the *associative law* and operators which obey it are called *associative*. Only associative operators are considered in the work here given.

For the repetition of an operator several times

$$ff = f^2, \quad fff = f^3, \quad f^m f^n = f^{m+n}, \tag{3}$$

the usual notation of powers is used. *The law of indices clearly holds;* for f^{m+n} means that f is applied $m+n$ times successively, whereas $f^m f^n$ means that it is applied n times and then m times more. Not applying the operator f at all would naturally be denoted by f^0, so that $f^0 u = u$ and the operator f^0 would be equivalent to multiplication by 1; the notation $f^0 = 1$ is adopted.

If for a given operation f there can be found an operation g such that the product $fg = f^0 = 1$ is equivalent to no operation, then g is called the *inverse* of f and notations such as

$$fg = 1, \quad g = f^{-1} = \frac{1}{f}, \quad ff^{-1} = f\frac{1}{f} = 1 \tag{4}$$

are regularly borrowed from arithmetic and algebra. Thus the inverse of the square is the square root, the inverse of sin is $\sin^{-1}$, the inverse of the logarithm is the exponential, the inverse of D is $\int$. Some operations have no inverse; multiplication by 0 is a case, and so is the square when applied to a negative number if only real numbers are considered. Other operations have more than one inverse; integration, the inverse of D, involves an arbitrary additive constant, and the inverse sine is a multiple valued function. It is therefore not always true that $f^{-1}f = 1$, but it is customary to mean by f^{-1} that particular inverse of f for which $f^{-1}f = ff^{-1} = 1$. Higher negative powers are defined by the equation $f^{-n} = (f^{-1})^n$, and it readily follows that $f^n f^{-n} = 1$, as may be seen by the example

$$f^3 f^{-3} = ff(f \cdot f^{-1})f^{-1}f^{-1} = f(f \cdot f^{-1})f^{-1} = ff^{-1} = 1.$$

The law of indices $f^m f^n = f^{m+n}$ also holds for negative indices, except in so far as $f^{-1}f$ may not be equal to 1 and may be required in the reduction of $f^m f^n$ to f^{m+n}.

If u, v, and $u+v$ are operands for the operator f and if

$$f(u+v) = fu + fv, \tag{5}$$

so that the operator applied to the sum gives the same result as the sum of the results of operating on each operand, then the operator f is called *linear* or *distributive.* If f denotes a function such that $f(x+y) = f(x) + f(y)$, it has been seen (Ex. 9, p. 45) that f must be equivalent to multiplication by a constant and $fx = Cx$. For a less specialized interpretation this is not so; for

$$D(u+v) = Du + Dv \quad \text{and} \quad \int(u+v) = \int u + \int v$$

are two of the fundamental formulas of calculus and show operators which are distributive and not equivalent to multiplication by a constant. Nevertheless it does follow by the same reasoning as used before (Ex. 9, p. 45), that $fnu = nfu$ if f is distributive and if n is a rational number.

Some operators have also the property of addition. Suppose that u is an operand and f, g are operators such that fu and gu are things that may be added together as $fu + gu$, then the *sum* of the operators, $f+g$, is defined by the equation $(f+g)u = fu + gu$. If furthermore the operators f, g, h are distributive, then

$$h(f+g) = hf + hg \quad \text{and} \quad (f+g)h = fh + gh, \tag{6}$$

and the multiplication of the operators becomes itself distributive. To prove this fact, it is merely necessary to consider that

$$h[(f+g)u] = h(fu + gu) = hfu + hgu$$

and

$$(f+g)(hu) = fhu + ghu.$$

Operators which are associative, commutative, distributive, and which admit addition may be treated algebraically, in so far as polynomials are concerned, by the ordinary algorisms of algebra; for it is by means of the associative, commutative, and distributive laws, and the law of indices that ordinary algebraic polynomials are rearranged, multiplied out, and factored. Now the operations of multiplication by constants and of differentiation or partial differentiation as applied to a function of one or more variables x, y, z, $\cdots$ do satisfy these laws. For instance

$$c(Du) = D(cu), \quad D_xD_yu = D_yD_xu, \quad (D_x + D_y)D_zu = D_xD_zu + D_yD_zu. \tag{7}$$

Hence, for example, if y be a function of x, the expression

$$D^ny + a_1D^{n-1}y + \cdots + a_{n-1}Dy + a_ny,$$

where the coefficients a are constants, may be written as

$$(D^n + a_1D^{n-1} + \cdots + a_{n-1}D + a_n)y \tag{8}$$

and may then be factored into the form

$$[(D-\alpha_1)(D-\alpha_2)\cdots(D-\alpha_{n-1})(D-\alpha_n)]y, \tag{8'}$$

where $\alpha_1, \alpha_2, \cdots, \alpha_n$ are the roots of the algebraic polynomial

$$x^n + a_1 x^{n-1} + \cdots + a_{n-1}x + a_n = 0.$$

EXERCISES

1. Show that $(fgh)^{-1} = h^{-1}g^{-1}f^{-1}$, that is, that the reciprocal of a product of operations is the product of the reciprocals in inverse order.

2. By definition the operator gfg^{-1} is called the transform of f by g. Show that (α) the transform of a product is the product of the transforms of the factors taken in the same order, and (β) the transform of the inverse is the inverse of the transform.

3. If $s \neq 1$ but $s^2 = 1$, the operator s is by definition said to be *involutory*. Show that (α) an involutory operator is equal to its own inverse; and conversely (β) if an operator and its inverse are equal, the operator is involutory; and (γ) if the product of two involutory operators is commutative, the product is itself involutory; and conversely (δ) if the product of two involutory operators is involutory, the operators are commutative.

4. If f and g are both distributive, so are the products fg and gf.

5. If f is distributive and n rational, show $fnu = nfu$.

6. Expand the following operators first by ordinary formal multiplication and second by applying the operators successively as indicated, and show the results are identical by translating both into familiar forms.

(α) $(D-1)(D-2)y$, *Ans.* $\dfrac{d^2y}{dx^2} - 3\dfrac{dy}{dx} + 2y$,

(β) $(D-1)D(D+1)y$, (γ) $D(D-2)(D+1)(D+3)y$.

7. Show that $(D-a)\left[e^{ax}\int e^{-ax}Xdx\right] = X$, where X is a function of x, and hence infer that $e^{ax}\int e^{-ax}(*)\,dx$ is the inverse of the operator $(D-a)(*)$.

8. Show that $D(e^{ax}y) = e^{ax}(D+a)y$ and hence generalize to show that if $P(D)$ denote any polynomial in D with constant coefficients, then

$$P(D)\cdot e^{ax}y = e^{ax}P(D+a)y.$$

Apply this to the following and check the results.

(α) $(D^2 - 3D + 2)e^{2x}y = e^{2x}(D^2 + D)y = e^{2x}\left(\dfrac{d^2y}{dx^2} + \dfrac{dy}{dx}\right)$,

(β) $(D^2 - 3D - 2)e^{x}y$, (γ) $(D^3 - 3D + 2)e^{x}y$.

9. If y is a function of x and $x = e^t$ show that

$$D_xy = e^{-t}D_ty,\ D_x^2y = e^{-2t}D_t(D_t-1)y,\ \cdots,\ D_x^py = e^{-pt}D_t(D_t-1)\cdots(D_t-p+1)y.$$

10. Is the expression $(hD_x + kD_y)^n$, which occurs in Taylor's Formula (§ 54), the nth power of the operator $hD_x + kD_y$ or is it merely a conventional symbol? The same question relative to $(xD_x + yD_y)^k$ occurring in Euler's Formula (§ 53)?

71. Complex numbers. In the formal solution of the equation $ax^2 + bx + c = 0$, where $b^2 < 4ac$, numbers of the form $m + n\sqrt{-1}$, where m and n are real, arise. Such numbers are called *complex* or *imaginary;* the part m is called the *real part* and $n\sqrt{-1}$ the *pure imaginary part* of the number. It is customary to write $\sqrt{-1} = i$ and to treat i as a literal quantity subject to the relation $i^2 = -1$. The definitions for the *equality, addition,* and *multiplication* of complex numbers are

$$\begin{aligned} a + bi = c + di \quad &\text{if and only if} \quad a = c,\ b = d, \\ [a + bi] + [c + di] &= (a + c) + (b + d)\,i, \\ [a + bi]\,[c + di] &= (ac - bd) + (ad + bc)\,i. \end{aligned} \tag{9}$$

It readily follows that *the commutative, associative, and distributive laws hold in the domain of complex numbers,* namely,

$$\begin{aligned} \alpha + \beta &= \beta + \alpha, & (\alpha + \beta) + \gamma &= \alpha + (\beta + \gamma), \\ \alpha\beta &= \beta\alpha, & (\alpha\beta)\,\gamma &= \alpha\,(\beta\gamma), \\ \alpha\,(\beta + \gamma) &= \alpha\beta + \alpha\gamma, & (\alpha + \beta)\,\gamma &= \alpha\gamma + \beta\gamma, \end{aligned} \tag{10}$$

where Greek letters have been used to denote complex numbers.

Division is accomplished by the method of rationalization.

$$\frac{a + bi}{c + di} = \frac{a + bi}{c + di}\,\frac{c - di}{c - di} = \frac{(ac + bd) + (bc - ad)\,i}{c^2 + d^2}. \tag{11}$$

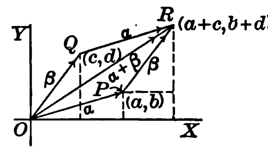
This is always possible except when $c^2 + d^2 = 0$, that is, when both c and d are 0. A complex number is defined as 0 when and only when its real and pure imaginary parts are both zero. With this definition 0 has the ordinary properties that $\alpha + 0 = \alpha$ and $\alpha 0 = 0$ and that $\alpha/0$ is impossible. Furthermore *if a product $\alpha\beta$ vanishes, either α or β vanishes.* For suppose

$$[a + bi]\,[c + di] = (ac - bd) + (ad + bc)\,i = 0.$$

Then
$$ac - bd = 0 \quad \text{and} \quad ad + bc = 0, \tag{12}$$

from which it follows that either $a = b = 0$ or $c = d = 0$. From the fact that a product cannot vanish unless one of its factors vanishes follow the ordinary laws of cancellation. In brief, *all the elementary laws of real algebra hold also for the algebra of complex numbers.*

By assuming a set of Cartesian coördinates in the xy-plane and associating the number $a + bi$ to the point (a, b), a *graphical representation* is obtained which is the counterpart of the number scale for real numbers. The point (a, b) alone or the directed line from the origin to the point (a, b) may be considered as representing the number $a + bi$. If OP and OQ are two directed lines representing the two numbers $a + bi$ and $c + di$, a reference to the figure shows that the line which represents the sum of the numbers is OR, the diagonal of the parallelogram of which OP and OQ are sides. Thus *the geometric law for adding complex numbers is the same as the law for compounding forces and is known as the parallelogram law.* A segment AB of a line possesses magnitude, the length AB, and direction, the direction of the line AB from A to B. *A quantity which has magnitude and direction is called a vector; and the parallelogram law is called the law of vector addition. Complex numbers* may therefore be regarded as *vectors.*



From the figure it also appears that OQ and PR have the same magnitude and direction, so that as vectors they are equal although they start from different points. As $OP + PR$ will be regarded as equal to $OP + OQ$, the definition of addition may be given as the triangle law instead of as the parallelogram law; namely, from the terminal end P of the first vector lay off the second vector PR and close the triangle by joining the initial end O of the first vector to the terminal end R of the second. The *absolute value* of a complex number $a + bi$ is the magnitude of its vector OP and is equal to $\sqrt{a^2 + b^2}$, the square root of the sum of the squares of its real part and of the coefficient of its pure imaginary part. The absolute value is denoted by $|a + bi|$ as in the case of reals. If α and β are two complex numbers, the rule $|\alpha| + |\beta| \geqq |\alpha + \beta|$ is a consequence of the fact that one side of a triangle is less than the sum of the other two. If the absolute value is given and the initial end of the vector is fixed, the terminal end is thereby constrained to lie upon a circle concentric with the initial end.

72. When the complex numbers are laid off from the origin, polar coördinates may be used in place of Cartesian. Then

$$r = \sqrt{a^2 + b^2}, \quad \phi = \tan^{-1} b/a^*, \quad a = r\cos\phi, \quad b = r\sin\phi$$
$$\text{and} \qquad a + ib = r(\cos\phi + i\sin\phi). \tag{13}$$

The absolute value r is often called the *modulus* or *magnitude* of the complex number; the angle ϕ is called the *angle* or *argument* of the number and suffers a certain indetermination in that $2n\pi$, where n is a positive or negative integer, may be added to ϕ without affecting the number. This polar representation is particularly useful in discussing products and quotients. For if

$$\alpha = r_1(\cos\phi_1 + i\sin\phi_1), \qquad \beta = r_2(\cos\phi_2 + i\sin\phi_2),$$
$$\text{then} \qquad \alpha\beta = r_1 r_2[\cos(\phi_1 + \phi_2) + i\sin(\phi_1 + \phi_2)], \tag{14}$$

* As both $\cos\phi$ and $\sin\phi$ are known, the quadrant of this angle is determined.

as may be seen by multiplication according to the rule. Hence the *magnitude of a product is the product of the magnitudes of the factors, and the angle of a product is the sum of the angles of the factors;* the general rule being proved by induction.

The interpretation of *multiplication by a complex number as an operation* is illuminating. Let β be the multiplicand and α the multiplier. As the product $\alpha\beta$ has a magnitude equal to the product of the magnitudes and an angle equal to the sum of the angles, the factor α used as a multiplier may be interpreted as effecting the rotation of β through the angle of α and the stretching of β in the ratio $|\alpha|:1$. From the geometric viewpoint, therefore, *multiplication by a complex number is an operation of rotation and stretching in the plane.* In the case of $\alpha = \cos\phi + i\sin\phi$ with $r = 1$, the operation is only of rotation and hence the factor $\cos\phi + i\sin\phi$ is often called a cyclic factor or versor. In particular the number $i = \sqrt{-1}$ will effect a rotation through 90° when used as a multiplier and is known as a quadrantal versor. The series of powers i, $i^2 = -1$, $i^3 = -i$, $i^4 = 1$ give rotations through 90°, 180°, 270°, 360°. This fact is often given as the reason for laying off pure imaginary numbers bi along an axis at right angles to the axis of reals.

As a particular product, the nth power of a complex number is

$$\alpha^n = (a + ib)^n = [r(\cos\phi + i\sin\phi)]^n = r^n(\cos n\phi + i\sin n\phi); \quad (15)$$

and

$$(\cos\phi + i\sin\phi)^n = \cos n\phi + i\sin n\phi, \quad (15')$$

which is a special case, is known as *De Moivre's Theorem* and is of use in evaluating the functions of $n\phi$; for the binomial theorem may be applied and the real and imaginary parts of the expansion may be equated to $\cos n\phi$ and $\sin n\phi$. Hence

$$\cos n\phi = \cos^n\phi - \frac{n(n-1)}{2!}\cos^{n-2}\phi\sin^2\phi + \frac{n(n-1)(n-2)(n-3)}{4!}\cos^{n-4}\phi\sin^4\phi - \cdots \quad (16)$$

$$\sin n\phi = n\cos^{n-1}\phi\sin\phi - \frac{n(n-1)(n-2)}{3!}\cos^{n-3}\phi\sin^3\phi + \cdots.$$

As the nth root $\sqrt[n]{\alpha}$ of α must be a number which when raised to the nth power gives α, the nth root may be written as

$$\sqrt[n]{\alpha} = \sqrt[n]{r}(\cos\phi/n + i\sin\phi/n). \quad (17)$$

The angle ϕ, however, may have any of the set of values

$$\phi, \quad \phi + 2\pi, \quad \phi + 4\pi, \quad \cdots, \quad \phi + 2(n-1)\pi,$$

and the nth parts of these give the n different angles

$$\frac{\phi}{n},\quad \frac{\phi}{n}+\frac{2\pi}{n},\quad \frac{\phi}{n}+\frac{4\pi}{n},\quad \cdots,\quad \frac{\phi}{n}+\frac{2(n-1)\pi}{n}. \tag{18}$$

Hence there may be found just n different nth roots of any given complex number (including, of course, the reals).

The *roots of unity* deserve mention. The equation $x^n = 1$ has in the real domain one or two roots according as n is odd or even. But if 1 be regarded as a complex number of which the pure imaginary part is zero, it may be represented by a point at a unit distance from the origin upon the axis of reals; the magnitude of 1 is 1 and the angle of 1 is $0, 2\pi, \cdots, 2(n-1)\pi$. The nth roots of 1 will therefore have the magnitude 1 and one of the angles $0, 2\pi/n, \cdots, 2(n-1)\pi/n$. The n nth roots are therefore

$$1,\quad \alpha = \cos\frac{2\pi}{n} + i\sin\frac{2\pi}{n},\quad \alpha^2 = \cos\frac{4\pi}{n} + i\sin\frac{4\pi}{n},\quad \cdots,$$
$$\alpha^{n-1} = \cos\frac{2(n-1)\pi}{n} + i\sin\frac{2(n-1)\pi}{n},$$

and may be evaluated with a table of natural functions. Now $x^n - 1 = 0$ is factorable as $(x-1)(x^{n-1} + x^{n-2} + \cdots + x + 1) = 0$, and it therefore follows that the nth roots other than 1 must all satisfy the equation formed by setting the second factor equal to 0. As α in particular satisfies this equation and the other roots are $\alpha^2, \cdots, \alpha^{n-1}$, it follows that the sum of the n nth roots of unity is zero.

EXERCISES

1. Prove the distributive law of multiplication for complex numbers.

2. By definition the pair of imaginaries $a + bi$ and $a - bi$ are called *conjugate imaginaries*. Prove that (α) the sum and the product of two conjugate imaginaries are real; and conversely (β) if the sum and the product of two imaginaries are both real, the imaginaries are conjugate.

3. Show that if $P(x, y)$ is a symmetric polynomial in x and y with real coefficients so that $P(x, y) = P(y, x)$, then if conjugate imaginaries be substituted for x and y, the value of the polynomial will be real.

4. Show that if $a + bi$ is a root of an algebraic equation $P(x) = 0$ with real coefficients, then $a - bi$ is also a root of the equation.

5. Carry out the indicated operations algebraically and make a graphical representation for every number concerned and for the answer:

$(\alpha)\ (1+i)^3,\qquad (\beta)\ (1+\sqrt{3}\,i)(1-i),\qquad (\gamma)\ (3+\sqrt{-2})(4+\sqrt{-5}),$

$(\delta)\ \dfrac{1+i}{1-i},\qquad (\epsilon)\ \dfrac{1+i\sqrt{3}}{1-i\sqrt{3}},\qquad (\zeta)\ \dfrac{5}{\sqrt{2}-i\sqrt{3}},$

$(\eta)\ \dfrac{(1-i)^2}{(1+i)^3},\qquad (\theta)\ \dfrac{1}{(1+i)^2}+\dfrac{1}{(1-i)^2},\qquad (\iota)\ \left(\dfrac{-1+\sqrt{-3}}{2}\right)^3.$

6. Plot and find the modulus and angle in the following cases:

$(\alpha)\ -2,\qquad (\beta)\ -2\sqrt{-1},\qquad (\gamma)\ 3+4i,\qquad (\delta)\ \tfrac{1}{2}-\tfrac{1}{2}\sqrt{-3}.$

7. Show that *the modulus of a quotient of two numbers is the quotient of the moduli and that the angle is the angle of the numerator less that of the denominator.*

8. Carry out the indicated operations trigonometrically and plot:

(α) The examples of Ex. 5, (β) $\sqrt{1+i}\sqrt{1-i}$, (γ) $\sqrt{-2+2\sqrt{3}i}$,
(δ) $(\sqrt{1+i}+\sqrt{1-i})^2$, ($\epsilon$) $\sqrt{\sqrt{2}+\sqrt{-2}}$, ($\zeta$) $\sqrt[3]{2+2\sqrt{3}i}$,
(η) $\sqrt[4]{16(\cos 200^\circ + i\sin 200^\circ)}$, ($\theta$) $\sqrt[5]{-1}$, (ι) $\sqrt[4]{8i}$.

9. Find the equations of analytic geometry which represent the transformation equivalent to multiplication by $\alpha = -1 + \sqrt{-3}$.

10. Show that $|z - \alpha| = r$, where z is a variable and α a fixed complex number, is the equation of the circle $(x-a)^2 + (y-b)^2 = r^2$.

11. Find $\cos 5x$ and $\cos 8x$ in terms of $\cos x$, and $\sin 6x$ and $\sin 7x$ in terms of $\sin x$.

12. Obtain to four decimal places the five roots $\sqrt[5]{1}$.

13. If $z = x + iy$ and $z' = x' + iy'$, show that $z' = (\cos\phi - i\sin\phi)z - \alpha$ is the formula for shifting the axes through the vector distance $\alpha = a + ib$ to the new origin (a, b) and turning them through the angle ϕ. Deduce the ordinary equations of transformation.

14. Show that $|z - \alpha| = k|z - \beta|$, where k is real, is the equation of a circle; specify the position of the circle carefully. Use the theorem: The locus of points whose distances to two fixed points are in a constant ratio is a circle the diameter of which is divided internally and externally in the same ratio by the fixed points.

15. The transformation $z' = \dfrac{az+b}{cz+d}$, where a, b, c, d are complex and $ad - bc \neq 0$, is called the *general linear transformation* of z into z'. Show that

$$|z' - \alpha'| = k|z' - \beta'| \quad \text{becomes} \quad |z - \alpha| = k\left|\frac{c\alpha' + d}{c\beta' + d}\right| \cdot |z - \beta|.$$

Hence infer that the transformation carries circles into circles, and points which divide a diameter internally and externally in the same ratio into points which divide some diameter of the new circle similarly, but generally with a different ratio.

73. Functions of a complex variable. Let $z = x + iy$ be a complex variable representable geometrically as a variable point in the xy-plane, which may be called the *complex plane*. As z determines the two real numbers x and y, any function $F(x, y)$ which is the sum of two single valued real functions in the form

$$F(x, y) = X(x, y) + iY(x, y) = R(\cos\Phi + i\sin\Phi) \tag{19}$$

will be completely determined in value if z is given. Such a function is called a *complex function* (and not a function of the complex variable, for reasons that will appear later). The magnitude and angle of the function are determined by

$$R = \sqrt{X^2 + Y^2}, \qquad \cos\Phi = \frac{X}{R}, \ \sin\Phi = \frac{Y}{R}. \tag{20}$$

The function F is continuous by definition when and only when both X and Y are continuous functions of (x, y); R is then continuous in (x, y) and F can vanish only when $R = 0$; the angle Φ regarded as a function of (x, y) is also continuous and determinate (except for the additive $2n\pi$) unless $R = 0$, in which case X and Y also vanish and the expression for Φ involves an indeterminate form in two variables and is generally neither determinate nor continuous (§ 44).

If the derivative of F with respect to z were sought for the value $z = a + ib$, the procedure would be entirely analogous to that in the case of a real function of a real variable. The increment $\Delta z = \Delta x + i\Delta y$ would be assumed for z and ΔF would be computed and the quotient $\Delta F/\Delta z$ would be formed. Thus by the Theorem of the Mean (§ 46),

$$\frac{\Delta F}{\Delta z} = \frac{\Delta X + i\Delta Y}{\Delta x + i\Delta y} = \frac{(X_x' + iY_x')\Delta x + (X_y' + iY_y')\Delta y}{\Delta x + i\Delta y} + \zeta, \tag{21}$$

where the derivatives are formed for (a, b) and where ζ is an infinitesimal complex number. When Δz approaches 0, both Δx and Δy must approach 0 without any implied relation between them. In general the limit of $\Delta F/\Delta z$ is a double limit (§ 44) and may therefore depend on the way in which Δx and Δy approach their limit 0.

Now if first $\Delta y \doteq 0$ and then subsequently $\Delta x \doteq 0$, the value of the limit of $\Delta F/\Delta z$ is $X_x' + iY_x'$ taken at the point (a, b); whereas if first $\Delta x \doteq 0$ and then $\Delta y \doteq 0$, the value is $-iX_y' + Y_y'$. Hence if the limit of $\Delta F/\Delta z$ is to be independent of the way in which Δz approaches 0, it is surely necessary that

$$\frac{\partial X}{\partial x} + i\frac{\partial Y}{\partial x} = -i\frac{\partial X}{\partial y} + \frac{\partial Y}{\partial y},$$

or
$$\frac{\partial X}{\partial x} = \frac{\partial Y}{\partial y} \quad \text{and} \quad \frac{\partial X}{\partial y} = -\frac{\partial Y}{\partial x}. \tag{22}$$

And conversely if these relations are satisfied, then

$$\frac{\Delta F}{\Delta z} = \left(\frac{\partial X}{\partial x} + i\frac{\partial Y}{\partial x}\right) + \zeta = \left(\frac{\partial Y}{\partial y} - i\frac{\partial X}{\partial y}\right) + \zeta;$$

and the limit is $X_x' + iY_x' = Y_y' - iX_y'$ taken at the point (a, b), and is independent of the way in which Δz approaches zero. The desirability of having at least the ordinary functions differentiable suggests the definition: *A complex function $F(x, y) = X(x, y) + iY(x, y)$ is considered as a function of the complex variable $z = x + iy$ when and only when X and Y are in general differentiable and satisfy the relations* (22). In this case *the derivative is*

$$F'(z) = \frac{dF}{dz} = \frac{\partial X}{\partial x} + i\frac{\partial Y}{\partial x} = \frac{\partial Y}{\partial y} - i\frac{\partial X}{\partial x}. \tag{23}$$

These conditions may also be expressed in polar coördinates (Ex. 2).

A few words about the function $\Phi(x, y)$. This is a multiple valued function of the variables (x, y), and the difference between two neighboring branches is the constant 2π. The application of the discussion of § 45 to this case shows at once that, in any simply connected region of the complex plane which contains no point (a, b) such that $R(a, b) = 0$, the different branches of $\Phi(x, y)$ may be entirely separated so that the value of Φ must return to its initial value when any closed curve is described by the point (x, y). If, however, the region is multiply connected or contains points for which $R = 0$ (which makes the region multiply connected because these points must be cut out), it may happen that there will be circuits for which Φ, although changing continuously, will not return to its initial value.. Indeed if it can be shown that Φ does not return to its initial value when changing continuously as (x, y) describes the boundary of a region simply connected except for the excised points, it may be inferred that there must be points in the region for which $R = 0$.

An application of these results may be made to give a very simple demonstration of *the fundamental theorem of algebra that every equation of the nth degree has at least one root.* Consider the function

$$F(z) = z^n + a_1 z^{n-1} + \cdots + a_{n-1}z + a_n = X(x, y) + iY(x, y), \quad = \sum_{k=0}^{n} a_k z^k, \quad a_0 = 1$$

where X and Y are found by writing z as $x + iy$ and expanding and rearranging. The functions X and Y will be polynomials in (x, y) and will therefore be everywhere finite and continuous in (x, y). Consider the angle Φ of F. Then

$$\Phi = \text{ang. of } F = \text{ang. of } z^n\left(1 + \frac{a_1}{z} + \cdots + \frac{a_{n-1}}{z^{n-1}} + \frac{a_n}{z^n}\right) = \text{ang. of } z^n + \text{ang. of } (1 + \cdots).$$

Next draw about the origin a circle of radius r so large that

$$\left|\frac{a_1}{z}\right| + \cdots + \left|\frac{a_{n-1}}{z^{n-1}}\right| + \left|\frac{a_n}{z^n}\right| = \frac{|a_1|}{r} + \cdots + \frac{|a_{n-1}|}{r^{n-1}} + \frac{|a_n|}{r^n} < \epsilon.$$

Then for all points z upon the circumference the angle of F is

$$\Phi = \text{ang. of } F = n(\text{ang. of } z) + \text{ang. of } (1 + \eta), \qquad |\eta| < \epsilon.$$

Now let the point (x, y) describe the circumference. The angle of z will change by 2π for the complete circuit. Hence Φ must change by $2n\pi$ and does not return to its initial value. Hence there is within the circle at least one point (a, b) for which $R(a, b) = 0$ and consequently for which $X(a, b) = 0$ and $Y(a, b) = 0$ and $F(a, b) = 0$. Thus if $\alpha = a + ib$, then $F(\alpha) = 0$ and the equation $F(z) = 0$ is seen to have at least the one root α. It follows that $z - \alpha$ is a factor of $F(z)$; and hence by induction it may be seen that $F(z) = 0$ has just n roots.

74. The discussion of the algebra of complex numbers showed how the sum, difference, product, quotient, real powers, and real roots of such numbers could be found, and hence made it possible to compute the value of any given algebraic expression or function of z for a given value of z. It remains to show that any algebraic expression in z is really a function of z in the sense that it has a derivative with respect to z, and to find the derivative. Now the differentiation of an algebraic function of the variable x was made to depend upon the formulas of differentiation, (6) and (7) of § 2. A glance at the methods of derivation of these formulas shows that they were proved by ordinary algebraic manipulations such as have been seen to be equally possible with imaginaries as with reals. It therefore may be concluded that *an algebraic expression in z has a derivative with respect to z and that derivative may be found just as if z were a real variable.*

The case of the elementary functions e^z, $\log z$, $\sin z$, $\cos z$, $\cdots$ other than algebraic is different; for these functions have not been defined for complex variables. Now in seeking to define these functions when z is complex, an effort should be made to define in such a way that: 1° when z is real, the new and the old definitions become identical; and 2° the rules of operation with the function shall be as nearly as possible the same for the complex domain as for the real. Thus it would be desirable that $De^z = e^z$ and $e^{z+w} = e^z e^w$, when z and w are complex. With these ideas in mind one may proceed to define the elementary functions for complex arguments. Let

$$e^z = R(x, y)[\cos \Phi(x, y) + i \sin \Phi(x, y)]. \tag{24}$$

The derivative of this function is, by the first rule of (23),

$$\begin{aligned} De^z &= \frac{\partial}{\partial x}(R \cos \Phi) + i \frac{\partial}{\partial x}(R \sin \Phi) \\ &= (R'_x \cos \Phi - R \sin \Phi \cdot \Phi'_x) + i(R'_x \sin \Phi + R \cos \Phi \cdot \Phi'_x), \end{aligned}$$

and if this is to be identical with e^z above, the equations

$$\begin{matrix} R'_x \cos \Phi - R\Phi'_x \sin \Phi = R \cos \Phi \\ R'_x \sin \Phi + R\Phi'_x \cos \Phi = R \sin \Phi \end{matrix} \quad \text{or} \quad \begin{matrix} R'_x = R \\ \Phi'_x = 0 \end{matrix}$$

must hold, where the second pair is obtained by solving the first. If the second form of the derivative in (23) had been used, the results would have been $R'_y = 0$, $\Phi'_y = 1$. It therefore appears that if the derivative of e^z, however computed, is to be e^z, then

$$R'_x = R, \quad R'_y = 0, \quad \Phi'_x = 0, \quad \Phi'_y = 1$$

are four conditions imposed upon R and Φ. These conditions will be satisfied if $R = e^x$ and $\Phi = y$.* Hence define

$$e^z = e^{x+iy} = e^x(\cos y + i \sin y). \tag{25}$$

* The use of the more general solutions $R = Ge^x$, $\Phi = y + C$ would lead to expressions which would not reduce to e^x when $y = 0$ and $z = x$ or would not satisfy $e^{z+w} = e^z e^w$.

With this definition De^z is surely e^z, and it is readily shown that the exponential law $e^{z+w} = e^z e^w$ holds.

For the special values $\frac{1}{2}\pi i$, πi, $2\pi i$ of z the value of e^z is

$$e^{\frac{1}{2}\pi i} = i, \quad e^{\pi i} = -1, \quad e^{2\pi i} = 1.$$

Hence it appears that if $2n\pi i$ be added to z, e^z is unchanged;

$$e^{z+2n\pi i} = e^z, \quad \text{period} \quad 2\pi i. \tag{26}$$

Thus *in the complex domain* e^z *has the period* $2\pi i$, just as $\cos x$ and $\sin x$ have the real period 2π. This relation is inherent; for

$$e^{yi} = \cos y + i \sin y, \quad e^{-yi} = \cos y - i \sin y,$$

and
$$\cos y = \frac{e^{yi} + e^{-yi}}{2}, \quad \sin y = \frac{e^{yi} - e^{-yi}}{2i}. \tag{27}$$

The trigonometric functions of a real variable y may be expressed in terms of the exponentials of yi and $-yi$. As the exponential has been defined for all complex values of z, it is natural to use (27) to define the trigonometric functions for complex values as

$$\cos z = \frac{e^{zi} + e^{-zi}}{2}, \quad \sin z = \frac{e^{zi} - e^{-zi}}{2i}. \tag{27'}$$

With these definitions the ordinary formulas for $\cos(z + w)$, $D \sin z$, $\cdots$ may be obtained and be seen to hold for complex arguments, just as the corresponding formulas were derived for the hyperbolic functions (§ 5).

As in the case of reals, the logarithm $\log z$ will be defined for complex numbers as the inverse of the exponential. Thus

$$\text{if} \quad e^z = w, \quad \text{then} \quad \log w = z + 2n\pi i, \tag{28}$$

where the periodicity of the function e^z shows that *the logarithm is not uniquely determined but admits the addition of* $2n\pi i$ *to any one of its values*, just as $\tan^{-1} x$ admits the addition of $n\pi$. If w is written as a complex number $u + iv$ with modulus $r = \sqrt{u^2 + v^2}$ and with the angle ϕ, it follows that

$$w = u + iv = r(\cos \phi + i \sin \phi) = re^{\phi i} = e^{\log r + \phi i}; \tag{29}$$

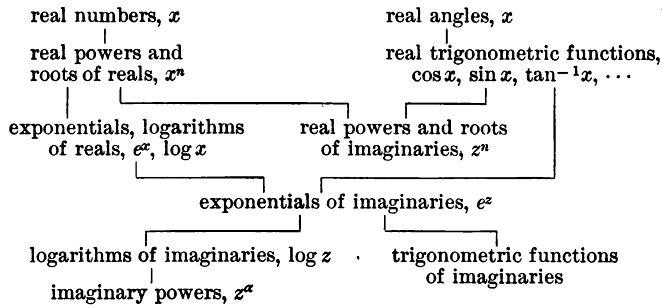
and
$$\log w = \log r + \phi i = \log \sqrt{u^2 + v^2} + i \tan^{-1}(v/u)$$

is the expression for the logarithm of w in terms of the modulus and angle of w; the $2n\pi i$ may be added if desired.

To this point the expression of a power a^b, where the exponent b is imaginary, has had no definition. The definition may now be given in terms of exponentials and logarithms. Let

$$a^b = e^{b \log a} \quad \text{or} \quad \log a^b = b \log a.$$

In this way the problem of computing a^b is reduced to one already solved. From the very definition it is seen that the logarithm of a power is the product of the exponent by the logarithm of the base, as in the case of reals. To indicate the path that has been followed in defining functions, a sort of family tree may be made.



EXERCISES

1. Show that the following complex functions satisfy the conditions (22) and are therefore functions of the complex variable z. Find $F'(z)$:

(α) $x^2 - y^2 + 2ixy$, (β) $x^3 - 3(xy^2 + x^2 - y^2) + i(3x^2y - y^3 - 6xy)$,

(γ) $\dfrac{x}{x^2+y^2} - i\dfrac{y}{x^2+y^2}$, ($\delta$) $\log\sqrt{x^2+y^2} + i\tan^{-1}\dfrac{y}{x}$,

(ϵ) $e^x\cos y + ie^x \sin y$, (ζ) $\sin x \sinh y + i\cos x \cosh y$.

2. Show that in polar coördinates the conditions for the existence of $F'(z)$ are

$$\frac{\partial X}{\partial r} = \frac{1}{r}\frac{\partial Y}{\partial \phi}, \quad \frac{\partial Y}{\partial r} = -\frac{1}{r}\frac{\partial X}{\partial \phi} \quad \text{with} \quad F'(z) = \left(\frac{\partial X}{\partial r} + i\frac{\partial Y}{\partial r}\right)(\cos\phi - i\sin\phi).$$

3. Use the conditions of Ex. 2 to show from $D\log z = z^{-1}$ that $\log z = \log r + \phi i$.

4. From the definitions given above prove the formulas

$$(\alpha)\ \sin(x+iy) = \sin x\cosh y + i\cos x \sinh y,$$
$$(\beta)\ \cos(x+iy) = \cos x\cosh y - i\sin x \sinh y,$$
$$(\gamma)\ \tan(x+iy) = \frac{\sin 2x + i\sinh 2y}{\cos 2x + \cosh 2y}.$$

5. Find to three decimals the complex numbers which express the values of:

(α) $e^{\frac{1}{4}\pi i}$, (β) e^i, (γ) $e^{\frac{1}{2}+\frac{1}{2}\sqrt{-3}}$, ($\delta$) e^{-1-i},

(ϵ) $\sin\frac{1}{4}\pi i$, (ζ) $\cos i$, (η) $\sin\left(\frac{1}{2}+\frac{1}{2}\sqrt{-3}\right)$, ($\theta$) $\tan(-1-i)$,

(ι) $\log(-1)$, (κ) $\log i$, (λ) $\log\left(\frac{1}{2}+\frac{1}{2}\sqrt{-3}\right)$, ($\mu$) $\log(-1-i)$.

6. Owing to the fact that $\log a$ is multiple valued, a^b is multiple valued in such a manner that any one value may be multiplied by $e^{2n\pi bi}$. Find one value of each of the following and several values of one of them:

(α) 2^i, (β) i^i, (γ) $\sqrt[i]{i}$, (δ) $\sqrt[i]{2}$, (ϵ) $\left(\frac{1}{2}+\frac{1}{2}\sqrt{-3}\right)^{\frac{3}{\pi}i+1}$.

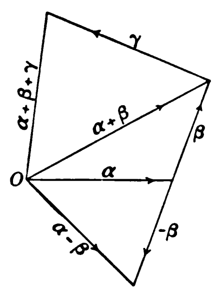
7. Show that $Da^z = a^z \log a$ when a and z are complex.

8. Show that $(a^b)^c = a^{bc}$; and fill in such other steps as may be suggested by the work in the text, which for the most part has merely been sketched in a broad way.

9. Show that if $f(z)$ and $g(z)$ are two functions of a complex variable, then $f(z) \pm g(z)$, $\alpha f(z)$ with α a complex constant, $f(z)g(z)$, $f(z)/g(z)$ are also functions of z.

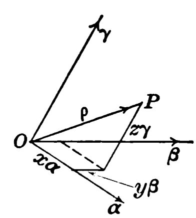
10. Obtain logarithmic expressions for the inverse trigonometric functions. Find $\sin^{-1} i$.

75. Vector sums and products. As stated in § 71, a vector is a quantity which has magnitude and direction. If the magnitudes of two vectors are equal and the directions of the two vectors are the same, the vectors are said to be equal irrespective of the position which they occupy in space. The vector $-\alpha$ is by definition a vector which has the same magnitude as α but the opposite direction. The vector $m\alpha$ is a vector which has the same direction as α (or the opposite) and is m (or $-m$) times as long. The law of vector or geometric addition is the parallelogram or triangle law (§ 71) and is still applicable when the vectors do not lie in a plane but have any directions in space; for any two vectors brought end to end determine a plane in which the construction may be carried out. Vectors will be designated by Greek small letters or by letters in heavy type. The relations of equality or similarity between triangles establish the rules



$$\alpha + \beta = \beta + \alpha, \quad \alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma, \quad m(\alpha + \beta) = m\alpha + m\beta \quad (30)$$

as true for vectors as well as for numbers whether real or complex. A vector is said to be zero when its magnitude is zero, and it is written 0. From the definition of addition it follows that $\alpha + 0 = \alpha$. In fact *as far as addition, subtraction, and multiplication by numbers are concerned, vectors obey the same formal laws as numbers.*



A vector ρ may be resolved into components parallel to any three given vectors α, β, γ which are not parallel to any one plane. For let a parallelepiped be constructed with its edges parallel to the three given vectors and with its diagonal equal to the vector whose components are desired. The edges of the parallelepiped are then certain multiples $x\alpha$, $y\beta$, $z\gamma$ of α, β, γ; and these are the desired components of ρ. The vector ρ may be written as

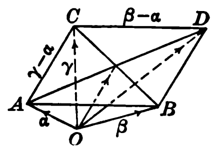
$$\rho = x\alpha + y\beta + z\gamma.* \tag{31}$$

It is clear that two equal vectors would necessarily have the same components along three given directions and that the components of a zero vector would all be zero. Just as the equality of two complex numbers involved the two equalities of the respective real and imaginary parts, so the equality of two vectors as

$$\rho = x\alpha + y\beta + z\gamma = x'\alpha + y'\beta + z'\gamma = \rho' \tag{31'}$$

involves the three equations $x = x'$, $y = y'$, $z = z'$.

As a problem in the use of vectors let there be given the three vectors α, β, γ from an assumed origin O to three vertices of a parallelogram; required the vector to the other vertex, the vector expressions for the sides and diagonals of the parallelogram, and the proof of the fact that the diagonals bisect each other. Consider the figure. The side AB is, by the triangle law, that vector which when added to $OA = \alpha$ gives $OB = \beta$, and hence it must be that $AB = \beta - \alpha$. In like manner $AC = \gamma - \alpha$. Now OD is the sum of OC and CD, and $CD = AB$; hence $OD = \gamma + \beta - \alpha$. The diagonal AD is the difference of the vectors OD and OA, and is therefore $\gamma + \beta - 2\alpha$. The diagonal BC is $\gamma - \beta$. Now the vector from O to the middle point of BC may be found by adding to OB one half of BC. Hence this vector is $\beta + \frac{1}{2}(\gamma - \beta)$ or $\frac{1}{2}(\beta + \gamma)$. In like manner the vector to the middle point of AD is seen to be $\alpha + \frac{1}{2}(\gamma + \beta - 2\alpha)$ or $\frac{1}{2}(\gamma + \beta)$, which is identical with the former. The two middle points therefore coincide and the diagonals bisect each other.



Let α and β be any two vectors, $|\alpha|$ and $|\beta|$ their respective lengths, and $\angle(\alpha, \beta)$ the angle between them. For convenience the vectors may be considered to be laid off from the same origin. The product of the lengths of the vectors by the cosine of the angle between the vectors is called the *scalar product*,

$$\text{scalar product} = \alpha\cdot\beta = |\alpha||\beta| \cos \angle(\alpha, \beta), \tag{32}$$

of the two vectors and is denoted by placing a dot between the letters. This combination, called the scalar product, is a number, not a vector. As $|\beta| \cos \angle(\alpha, \beta)$ is the projection of β upon the direction of α, the scalar product may be stated to be equal to the product of the length of either vector by the length of the projection of the other upon it. In particular if either vector were of unit length, the scalar product would be the projection of the other upon it, with proper regard for

* The numbers x, y, z are the oblique coördinates of the terminal end of ρ (if the initial end be at the origin) referred to a set of axes which are parallel to α, β, γ and upon which the unit lengths are taken as the lengths of α, β, γ respectively.

the sign; and if both vectors are unit vectors, the product is the cosine of the angle between them.

The scalar product, from its definition, is *commutative* so that $\alpha \cdot \beta = \beta \cdot \alpha$. Moreover $(m\alpha)\cdot\beta = \alpha\cdot(m\beta) = m\,(\alpha\cdot\beta)$, thus allowing a numerical factor m to be combined with either factor of the product. Furthermore the *distributive law*

$$\alpha\cdot(\beta + \gamma) = \alpha\cdot\beta + \alpha\cdot\gamma \quad \text{or} \quad (\alpha + \beta)\cdot\gamma = \alpha\cdot\gamma + \beta\cdot\gamma \tag{33}$$

is satisfied as in the case of numbers. For if α be written as the product $a\alpha_1$ of its length a by a vector α_1 of unit length in the direction of α, the first equation becomes

$$a\alpha_1\cdot(\beta + \gamma) = a\alpha_1\cdot\beta + a\alpha_1\cdot\gamma \quad \text{or} \quad \alpha_1\cdot(\beta + \gamma) = \alpha_1\cdot\beta + \alpha_1\cdot\gamma.$$

And now $\alpha_1\cdot(\beta + \gamma)$ is the projection of the sum $\beta + \gamma$ upon the direction of α, and $\alpha_1\cdot\beta + \alpha_1\cdot\gamma$ is the sum of the projections of β and γ upon this direction; by the law of projections these are equal and hence the distributive law is proved.

The associative law does not hold for scalar products; for $(\alpha\cdot\beta)\,\gamma$ means that the vector γ is multiplied by the number $\alpha\cdot\beta$, whereas $\alpha\,(\beta\cdot\gamma)$ means that α is multiplied by $(\beta\cdot\gamma)$, a very different matter. The laws of cancellation cannot hold; for if

$$\alpha\cdot\beta = 0, \quad \text{then} \quad |\alpha||\beta|\cos\angle(\alpha, \beta) = 0, \tag{34}$$

and the vanishing of the scalar product $\alpha\cdot\beta$ implies either that one of the factors is 0 or that the two vectors are perpendicular. In fact $\alpha\cdot\beta = 0$ is called the *condition of perpendicularity*. It should be noted, however, that if a vector ρ satisfies

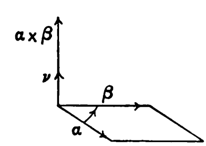
$$\rho\cdot\alpha = 0, \qquad \rho\cdot\beta = 0, \qquad \rho\cdot\gamma = 0, \tag{35}$$

three conditions of perpendicularity with three vectors α, β, γ not parallel to the same plane, the inference is that $\rho = 0$.

76. Another product of two vectors is the *vector product*,

$$\text{vector product} = \alpha\times\beta = \nu|\alpha||\beta|\sin\angle(\alpha, \beta), \tag{36}$$

where ν represents a vector of unit length normal to the plane of α and β upon that side on which rotation from α to β through an angle of less than 180° appears positive or counterclockwise. Thus the vector product is itself a vector of which the direction is perpendicular to each factor, and of which the magnitude is the product of the magnitudes into the sine of the included angle. The magnitude is therefore equal to the area of the parallelogram of which the vectors α and β are the sides.



The vector product will be represented by a cross inserted between the letters.

As rotation from β to α is the opposite of that from α to β, it follows from the definition of the vector product that

$$\beta\times\alpha = -\alpha\times\beta, \quad \text{not} \quad \alpha\times\beta = \beta\times\alpha, \tag{37}$$

and the product is *not commutative*, the order of the factors must be carefully observed. Furthermore the equation

$$\alpha\times\beta = \nu|\alpha||\beta|\sin\angle(\alpha, \beta) = 0 \tag{38}$$

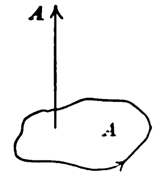
implies either that one of the factors vanishes or that the vectors α and β are parallel. Indeed the condition $\alpha\times\beta = 0$ is called the *condition of parallelism*. The laws of cancellation do not hold. The associative law also does not hold; for $(\alpha\times\beta)\times\gamma$ is a vector perpendicular to $\alpha\times\beta$ and γ, and since $\alpha\times\beta$ is perpendicular to the plane of α and β, the vector $(\alpha\times\beta)\times\gamma$ perpendicular to it must lie in the plane of α and β; whereas the vector $\alpha\times(\beta\times\gamma)$, by similar reasoning, must lie in the plane of β and γ; and hence the two vectors cannot be equal except in the very special case where each was parallel to β which is common to the two planes.

But the operation $(m\alpha)\times\beta = \alpha\times(m\beta) = m(\alpha\times\beta)$, which consists in allowing the transference of a numerical factor to any position in the product, does hold; and so does the *distributive law*

$$\alpha\times(\beta+\gamma) = \alpha\times\beta + \alpha\times\gamma \quad \text{and} \quad (\alpha+\beta)\times\gamma = \alpha\times\gamma + \beta\times\gamma, \tag{39}$$

the proof of which will be given below. In expanding according to the distributive law care must be exercised to keep the order of the factors in each vector product the same on both sides of the equation, owing to the failure of the commutative law; an interchange of the order of the factors changes the sign. It might seem as if any algebraic operations where so many of the laws of elementary algebra fail as in the case of vector products would be too restricted to be very useful; that this is not so is due to the astonishingly great number of problems in which the analysis can be carried on with only the laws of addition and the distributive law of multiplication combined with the possibility of transferring a numerical factor from one position to another in a product; in addition to these laws, the scalar product $\alpha\cdot\beta$ is commutative and the vector product $\alpha\times\beta$ is commutative except for change of sign.

In addition to segments of lines, *plane areas may be regarded as vector quantities;* for a plane area has magnitude (the amount of the area) and direction (the direction of the normal to its plane). To specify on which side of the plane the normal lies, some convention must be made. If the area is part of a surface inclosing a portion of space, the normal is taken as the exterior normal. If the area lies in an isolated plane, its positive side is determined only in connection with some assigned direction of description of its bounding curve; the rule is: If a person is assumed to walk along the boundary of an area in an assigned direction and upon that side of the plane which causes the inclosed area to lie upon his left, he is said to be upon the positive side (for the assigned direction of description of the boundary), and the vector which represents the area is the normal to that side. It has been mentioned that the vector product represented an area.



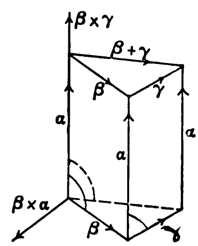
That the projection of a plane area upon a given plane gives an area which is the original area multiplied by the cosine of the angle between the two planes is a fundamental fact of projection, following from the simple fact that lines parallel to the intersection of the two planes are unchanged in length whereas lines perpendicular to the intersection are multiplied by the cosine of the angle between the planes. As the angle between the normals is the same as that between the planes, *the projection of an area upon a plane and the projection of the vector representing the area upon the normal to the plane are equivalent.* The projection of a *closed* area upon a plane is zero; for the area in the projection is covered twice (or an even number of times) with opposite signs and the total algebraic sum is therefore 0.

To prove the law $\alpha\times(\beta+\gamma)=\alpha\times\beta+\alpha\times\gamma$ and illustrate the use of the vector interpretation of areas, construct a triangular prism with the triangle on β, γ, and $\beta+\gamma$ as base and α as lateral edge. The total vector expression for the surface of this prism is

$$\beta\times\alpha+\gamma\times\alpha+\alpha\times(\beta+\gamma)+\tfrac{1}{2}(\beta\times\gamma)-\tfrac{1}{2}\beta\times\gamma=0,$$

and vanishes because the surface is closed. A cancellation of the equal and opposite terms (the two bases) and a simple transposition combined with the rule $\beta\times\alpha=-\alpha\times\beta$ gives the result

$$\alpha\times(\beta+\gamma)=-\beta\times\alpha-\gamma\times\alpha=\alpha\times\beta+\alpha\times\gamma.$$



A system of *vectors of reference* which is particularly useful consists of three vectors **i**, **j**, **k** of unit length directed along the axes X, Y, Z drawn so that rotation from X to Y appears positive from the side of the xy-plane upon which Z lies. The components of any vector **r** drawn from the origin to the point (x, y, z) are

$$x\mathbf{i},\quad y\mathbf{j},\quad z\mathbf{k},\quad \text{and}\quad \mathbf{r}=x\mathbf{i}+y\mathbf{j}+z\mathbf{k}.$$

The products of **i**, **j**, **k** into each other are, from the definitions,

$$\begin{gathered}\mathbf{i}\cdot\mathbf{i} = \mathbf{j}\cdot\mathbf{j} = \mathbf{k}\cdot\mathbf{k} = 1,\\ \mathbf{i}\cdot\mathbf{j} = \mathbf{j}\cdot\mathbf{i} = \mathbf{j}\cdot\mathbf{k} = \mathbf{k}\cdot\mathbf{j} = \mathbf{k}\cdot\mathbf{i} = \mathbf{i}\cdot\mathbf{k} = 0,\\ \mathbf{i}\times\mathbf{i} = \mathbf{j}\times\mathbf{j} = \mathbf{k}\times\mathbf{k} = 0,\\ \mathbf{i}\times\mathbf{j} = -\mathbf{j}\times\mathbf{i} = \mathbf{k},\quad \mathbf{j}\times\mathbf{k} = -\mathbf{k}\times\mathbf{j} = \mathbf{i},\quad \mathbf{k}\times\mathbf{i} = -\mathbf{i}\times\mathbf{k} = \mathbf{j}.\end{gathered} \tag{40}$$

By means of these products and the distributive laws for scalar and vector products, any given products may be expanded. Thus if

$$\alpha = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k} \quad\text{and}\quad \beta = b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k},$$

then

$$\begin{gathered}\alpha\cdot\beta = a_1b_1 + a_2b_2 + a_3b_3,\\ \alpha\times\beta = (a_2b_3 - a_3b_2)\mathbf{i} + (a_3b_1 - a_1b_3)\mathbf{j} + (a_1b_2 - a_2b_1)\mathbf{k},\end{gathered} \tag{41}$$

by direct multiplication. In this way a passage may be made from vector formulas to Cartesian formulas whenever desired.

EXERCISES

1. Prove geometrically that $\alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma$ and $m(\alpha + \beta) = m\alpha + m\beta$.

2. If α and β are the vectors from an assumed origin to A and B and if C divides AB in the ratio $m : n$, show that the vector to C is $\gamma = (n\alpha + m\beta)/(m + n)$.

3. In the parallelogram $ABCD$ show that the line BE connecting the vertex to the middle point of the opposite side CD is trisected by the diagonal AD and trisects it.

4. Show that the medians of a triangle meet in a point and are trisected.

5. If m_1 and m_2 are two masses situated at P_1 and P_2, the *center of gravity* or *center of mass* of m_1 and m_2 is defined as that point G on the line P_1P_2 which divides P_1P_2 inversely as the masses. Moreover if G_1 is the center of mass of a number of masses of which the total mass is M_1 and if G_2 is the center of mass of a number of other masses whose total mass is M_2, the same rule applied to M_1 and M_2 and G_1 and G_2 gives the center of gravity G of the total number of masses. Show that

$$\bar{\mathbf{r}} = \frac{m_1\mathbf{r}_1 + m_2\mathbf{r}_2}{m_1 + m_2} \quad\text{and}\quad \bar{\mathbf{r}} = \frac{m_1\mathbf{r}_1 + m_2\mathbf{r}_2 + \cdots + m_n\mathbf{r}_n}{m_1 + m_2 + \cdots + m_n} = \frac{\Sigma m\mathbf{r}}{\Sigma m},$$

where $\bar{\mathbf{r}}$ denotes the vector to the center of gravity. Resolve into components to show

$$\bar{x} = \frac{\Sigma mx}{\Sigma m},\quad \bar{y} = \frac{\Sigma my}{\Sigma m},\quad \bar{z} = \frac{\Sigma mz}{\Sigma m}.$$

6. If α and β are two fixed vectors and ρ a variable vector, all being laid off from the same origin, show that $(\rho - \beta)\cdot\alpha = 0$ is the equation of a plane through the end of β perpendicular to α.

7. Let α, β, γ be the vectors to the vertices A, B, C of a triangle. Write the three equations of the planes through the vertices perpendicular to the opposite sides. Show that the third of these can be derived as a combination of the other two; and hence infer that the three planes have a line in common and that the perpendiculars from the vertices of a triangle meet in a point.

8. Solve the problem analogous to Ex. 7 for the perpendicular bisectors of the sides.

9. Note that the length of a vector is $\sqrt{\alpha\cdot\alpha}$. If α, β, and $\gamma = \beta - \alpha$ are the three sides of a triangle, expand $\gamma\cdot\gamma = (\beta - \alpha)\cdot(\beta - \alpha)$ to obtain the law of cosines.

10. Show that the sum of the squares of the diagonals of a parallelogram equals the sum of the squares of the sides. What does the difference of the squares of the diagonals equal?

11. Show that $\dfrac{\alpha\cdot\beta}{\alpha\cdot\alpha}\alpha$ and $\dfrac{(\alpha\times\beta)\times\alpha}{\alpha\cdot\alpha}$ are the components of β parallel and perpendicular to α by showing 1° that these vectors have the right direction, and 2° that they have the right magnitude.

12. If α, β, γ are the three edges of a parallelepiped which start from the same vertex, show that $(\alpha\times\beta)\cdot\gamma$ is the volume of the parallelepiped, the volume being considered positive if γ lies on the same side of the plane of α and β with the vector $\alpha\times\beta$.

13. Show by Ex. 12 that $(\alpha\times\beta)\cdot\gamma = \alpha\cdot(\beta\times\gamma)$ and $(\alpha\times\beta)\cdot\gamma = (\beta\times\gamma)\cdot\alpha$; and hence infer that in a product of three vectors with cross and dot, the position of the cross and dot may be interchanged and the order of the factors may be permuted cyclically without altering the value. Show that the vanishing of $(\alpha\times\beta)\cdot\gamma$ or any of its equivalent expressions denotes that α, β, γ are parallel to the same plane; the condition $\alpha\times\beta\cdot\gamma = 0$ is called the condition of complanarity.

14. Assuming $\alpha = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}$, $\beta = b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k}$, $\gamma = c_1\mathbf{i} + c_2\mathbf{j} + c_3\mathbf{k}$, expand $\alpha\cdot\gamma$, $\alpha\cdot\beta$, and $\alpha\times(\beta\times\gamma)$ in terms of the coefficients to show

$$\alpha\times(\beta\times\gamma) = (\alpha\cdot\gamma)\beta - (\alpha\cdot\beta)\gamma; \quad \text{and hence} \quad (\alpha\times\beta)\times\gamma = (\alpha\cdot\gamma)\beta - (\gamma\cdot\beta)\alpha.$$

15. The formulas of Ex. 14 for expanding a product with two crosses and the rule of Ex. 13 that a dot and a cross may be interchanged may be applied to expand

$$(\alpha\times\beta)\times(\gamma\times\delta) = (\alpha\cdot\gamma\times\delta)\beta - (\beta\cdot\gamma\times\delta)\alpha = (\alpha\times\beta\cdot\delta)\gamma - (\alpha\times\beta\cdot\gamma)\delta$$

and
$$(\alpha\times\beta)\cdot(\gamma\times\delta) = (\alpha\cdot\gamma)(\beta\cdot\delta) - (\beta\cdot\gamma)(\alpha\cdot\delta).$$

16. If α and β are two unit vectors in the xy-plane inclined at angles θ and ϕ to the x-axis, show that

$$\alpha = \mathbf{i}\cos\theta + \mathbf{j}\sin\theta, \quad \beta = \mathbf{i}\cos\phi + \mathbf{j}\sin\phi;$$

and from the fact that $\alpha\cdot\beta = \cos(\phi - \theta)$ and $\alpha\times\beta = \mathbf{k}\sin(\phi - \theta)$ obtain by multiplication the trigonometric formulas for $\sin(\phi - \theta)$ and $\cos(\phi - \theta)$.

17. If l, m, n are direction cosines, the vector $l\mathbf{i} + m\mathbf{j} + n\mathbf{k}$ is a vector of unit length in the direction for which l, m, n are direction cosines. Show that the condition for perpendicularity of two directions (l, m, n) and (l', m', n') is $ll' + mm' + nn' = 0$.

18. With the same notations as in Ex. 14 show that

$$\alpha\cdot\alpha = a_1^2 + a_2^2 + a_3^2 \quad \text{and} \quad \alpha\times\beta = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix} \quad \text{and} \quad \alpha\times\beta\cdot\gamma = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}.$$

19. Compute the scalar and vector products of these pairs of vectors:

$$(\alpha)\begin{cases}6\mathbf{i}+0.3\mathbf{j}-5\mathbf{k}\\0.1\mathbf{i}-4.2\mathbf{j}+2.5\mathbf{k},\end{cases}\quad(\beta)\begin{cases}\mathbf{i}+2\mathbf{j}+3\mathbf{k}\\-3\mathbf{i}-2\mathbf{j}+\mathbf{k},\end{cases}\quad(\gamma)\begin{cases}\mathbf{i}+\mathbf{k}\\\mathbf{j}+\mathbf{i}.\end{cases}$$

20. Find the areas of the parallelograms defined by the pairs of vectors in Ex. 19. Find also the sine and cosine of the angles between the vectors.

21. Prove $\alpha\times[\beta\times(\gamma\times\delta)]=(\alpha\cdot\gamma\times\delta)\beta-\alpha\cdot\beta\,\gamma\times\delta=\beta\cdot\delta\,\alpha\times\gamma-\beta\cdot\gamma\,\alpha\times\delta$.

22. What is the area of the triangle (1, 1, 1), (0, 2, 3), (0, 0, − 1)?

77. Vector differentiation. As the fundamental rules of differentiation depend on the laws of subtraction, multiplication by a number, the distributive law, and the rules permitting rearrangement, it follows that the rules must be applicable to expressions containing vectors without any changes except those implied by the fact that $\alpha\times\beta\neq\beta\times\alpha$. As an illustration consider the application of the definition of differentiation to the vector product $\mathbf{u}\times\mathbf{v}$ of two vectors which are supposed to be functions of a numerical variable, say x. Then

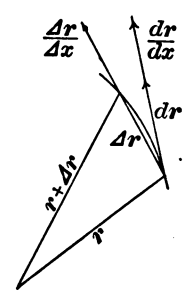
$$\Delta(\mathbf{u}\times\mathbf{v})=(\mathbf{u}+\Delta\mathbf{u})\times(\mathbf{v}+\Delta\mathbf{v})-\mathbf{u}\times\mathbf{v}$$
$$=\mathbf{u}\times\Delta\mathbf{v}+\Delta\mathbf{u}\times\mathbf{v}+\Delta\mathbf{u}\times\Delta\mathbf{v},$$

$$\frac{\Delta(\mathbf{u}\times\mathbf{v})}{\Delta x}=\mathbf{u}\times\frac{\Delta\mathbf{v}}{\Delta x}+\frac{\Delta\mathbf{u}}{\Delta x}\times\mathbf{v}+\frac{\Delta\mathbf{u}\times\Delta\mathbf{v}}{\Delta x},$$

$$\frac{d(\mathbf{u}\times\mathbf{v})}{dx}=\lim_{\Delta x\doteq 0}\frac{\Delta(\mathbf{u}\times\mathbf{v})}{\Delta x}=\mathbf{u}\times\frac{d\mathbf{v}}{dx}+\frac{d\mathbf{u}}{dx}\times\mathbf{v}.$$

Here the ordinary rule for a product is seen to hold, except that *the order of the factors must not be interchanged.*

The interpretation of the derivative is important. Let the variable vector $\mathbf{r}$ be regarded as a function of some variable, say x, and suppose $\mathbf{r}$ is laid off from an assumed origin so that, as x varies, the terminal point of $\mathbf{r}$ describes a curve. The increment $\Delta\mathbf{r}$ of $\mathbf{r}$ corresponding to Δx is a vector quantity and in fact is the chord of the curve as indicated. *The derivative*



$$\frac{d\mathbf{r}}{dx}=\lim\frac{\Delta\mathbf{r}}{\Delta x},\quad\frac{d\mathbf{r}}{ds}=\lim\frac{\Delta\mathbf{r}}{\Delta s}=\mathbf{t}\qquad(42)$$

is therefore a vector tangent to the curve; in particular if the variable x were the arc s, the derivative would have the magnitude unity and would be a unit vector tangent to the curve.

The derivative or differential of a vector of constant length is perpendicular to the vector. This follows from the fact that the vector then describes a circle concentric with the origin. It may also be seen analytically from the equation

$$d(\mathbf{r}\cdot\mathbf{r}) = d\mathbf{r}\cdot\mathbf{r} + \mathbf{r}\cdot d\mathbf{r} = 2\,\mathbf{r}\cdot d\mathbf{r} = d\,\text{const.} = 0. \tag{43}$$

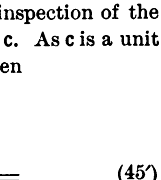
If the vector of constant length is of length unity, the increment $\Delta\mathbf{r}$ is the chord in a unit circle and, apart from infinitesimals of higher order, it is equal in magnitude to the angle subtended at the center. Consider then the derivative of the unit tangent $\mathbf{t}$ to a curve with respect to the arc s. The magnitude of $d\mathbf{t}$ is the angle the tangent turns through and the direction of $d\mathbf{t}$ is normal to $\mathbf{t}$ and hence to the curve. The vector quantity,

$$\text{curvature} \quad \mathbf{C} = \frac{d\mathbf{t}}{ds} = \frac{d^2\mathbf{r}}{ds^2}, \tag{44}$$

therefore has the magnitude of the curvature (by the definition in § 42) and the direction of the interior normal to the curve.

This work holds equally for plane or space curves. In the case of a space curve the plane which contains the tangent $\mathbf{t}$ and the curvature $\mathbf{C}$ is called the osculating plane (§ 41). By definition (§ 42) the *torsion of a space curve* is the rate of turning of the osculating plane with the arc, that is, $d\psi/ds$. To find the torsion by vector methods let $\mathbf{c}$ be a unit vector $\mathbf{C}/\sqrt{\mathbf{C}\cdot\mathbf{C}}$ along $\mathbf{C}$. Then as $\mathbf{t}$ and $\mathbf{c}$ are perpendicular, $\mathbf{n} = \mathbf{t}\times\mathbf{c}$ is a unit vector perpendicular to the osculating plane and $d\mathbf{n}$ will equal $d\psi$ in magnitude. Hence as a vector quantity the torsion is

$$\mathbf{T} = \frac{d\mathbf{n}}{ds} = \frac{d(\mathbf{t}\times\mathbf{c})}{ds} = \frac{d\mathbf{t}}{ds}\times\mathbf{c} + \mathbf{t}\times\frac{d\mathbf{c}}{ds} = \mathbf{t}\times\frac{d\mathbf{c}}{ds}, \tag{45}$$

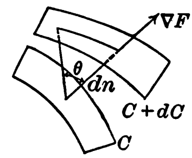


where (since $d\mathbf{t}/ds = \mathbf{C}$, and $\mathbf{c}$ is parallel to $\mathbf{C}$) the first term drops out. Next note that $d\mathbf{n}$ is perpendicular to $\mathbf{n}$ because it is the differential of a unit vector, and is perpendicular to $\mathbf{t}$ because $d\mathbf{n} = d(\mathbf{t}\times\mathbf{c}) = \mathbf{t}\times d\mathbf{c}$ and $\mathbf{t}\cdot(\mathbf{t}\times d\mathbf{c}) = 0$ since $\mathbf{t}$, $\mathbf{t}$, $d\mathbf{c}$ are necessarily complanar (Ex. 12, p. 169). Hence $\mathbf{T}$ is parallel to $\mathbf{c}$. It is convenient to consider the torsion as positive when the osculating plane seems to turn in the positive direction when viewed from the side of the normal plane upon which $\mathbf{t}$ lies. An inspection of the figure shows that in this case $d\mathbf{n}$ has the direction $-\mathbf{c}$ and not $+\mathbf{c}$. As $\mathbf{c}$ is a unit vector, the numerical value of the torsion is therefore $-\mathbf{c}\cdot\mathbf{T}$. Then

$$\begin{aligned} T &= -\mathbf{c}\cdot\mathbf{T} = -\,\mathbf{c}\cdot\mathbf{t}\times\frac{d\mathbf{c}}{ds} = -\,\mathbf{c}\cdot\mathbf{t}\times\frac{d}{ds}\frac{\mathbf{C}}{\sqrt{\mathbf{C}\cdot\mathbf{C}}} \\ &= -\,\mathbf{c}\cdot\mathbf{t}\times\left[\frac{d^3\mathbf{r}}{ds^3}\frac{1}{\sqrt{\mathbf{C}\cdot\mathbf{C}}} + \mathbf{C}\frac{d}{ds}\frac{1}{\sqrt{\mathbf{C}\cdot\mathbf{C}}}\right] = -\mathbf{c}\cdot\mathbf{t}\times\frac{d^3\mathbf{r}}{ds^3}\frac{1}{\sqrt{\mathbf{C}\cdot\mathbf{C}}} \\ &= \mathbf{t}\cdot\frac{\mathbf{C}}{\mathbf{C}\cdot\mathbf{C}}\times\frac{d^3\mathbf{r}}{ds^3} = \frac{\mathbf{r}'\cdot\mathbf{r}''\times\mathbf{r}'''}{\mathbf{r}''\cdot\mathbf{r}''}, \end{aligned} \tag{45'}$$

where differentiation with respect to s is denoted by accents.

78. Another sort of relation between vectors and differentiation comes to light in connection with the normal and directional derivatives (§ 48). If $F(x, y, z)$ is a function which has a definite value at each point of space and if the two neighboring surfaces $F = C$ and $F = C + dC$ are considered, the normal derivative of F is the rate of change of F along the normal to the surfaces and is written dF/dn. The rate of change of F along the normal to the surface $F = C$ is more rapid than along any other direction; for the change in F between the two surfaces is $dF = dC$ and is constant, whereas the distance dn between the two surfaces is least (apart from infinitesimals of higher order) along the normal. In fact if dr denote the distance along any other direction, the relations shown by the figure are



$$dr = \sec\theta dn \quad \text{and} \quad \frac{dF}{dr} = \frac{dF}{dn}\cos\theta. \tag{46}$$

If now $\mathbf{n}$ denote a vector of unit length normal to the surface, *the product* $\mathbf{n}\,dF/dn$ *will be a vector quantity which has both the magnitude and the direction of most rapid increase of* F. Let

$$\mathbf{n}\frac{dF}{dn} = \nabla F = \text{grad}\, F \tag{47}$$

be the symbolic expressions for this vector, where ∇F is read as "del F" and grad F is read as "the gradient of F." If $d\mathbf{r}$ be the vector of which dr is the length, the scalar product $\mathbf{n}\cdot d\mathbf{r}$ is precisely $\cos\theta dr$, and hence it follows that

$$d\mathbf{r}\cdot\nabla F = dF \quad \text{and} \quad \mathbf{r}_1\cdot\nabla F = \frac{dF}{dr}, \tag{48}$$

where $\mathbf{r}_1$ is a unit vector in the direction $d\mathbf{r}$. The second of the equations shows that *the directional derivative in any direction is the component or projection of the gradient in that direction.*

From this fact the expression of the gradient may be found in terms of its components along the axes. For the derivatives of F along the axes are $\partial F/\partial x$, $\partial F/\partial y$, $\partial F/\partial z$, and as these are the components of ∇F along the directions $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$, the result is

$$\nabla F = \text{grad}\, F = \mathbf{i}\frac{\partial F}{\partial x} + \mathbf{j}\frac{\partial F}{\partial y} + \mathbf{k}\frac{\partial F}{\partial z}.$$

Hence

$$\nabla = \mathbf{i}\frac{\partial}{\partial x} + \mathbf{j}\frac{\partial}{\partial y} + \mathbf{k}\frac{\partial}{\partial z} \tag{49}$$

may be regarded as a symbolic vector-differentiating operator which when applied to F gives the gradient of F. The product

$$d\mathbf{r}\cdot\nabla F = \left(dx\frac{\partial}{\partial x} + dy\frac{\partial}{\partial y} + dz\frac{\partial}{\partial z}\right)F = dF \tag{50}$$

is immediately seen to give the ordinary expression for dF. From this form of grad F it does not appear that the gradient of a function is independent of the choice of axes, but from the manner of derivation of ∇F first given it does appear that grad F is a definite vector quantity independent of the choice of axes.

In the case of any given function F the gradient may be found by the application of the formula (49); but in many instances it may also be found by means of the important relation $d\mathbf{r}\cdot\nabla F = dF$ of (48). For instance to prove the formula $\nabla(FG) = F\nabla G + G\nabla F$, the relation may be applied as follows:

$$\begin{aligned} d\mathbf{r}\cdot\nabla(FG) &= d(FG) = FdG + GdF \\ &= Fd\mathbf{r}\cdot\nabla G + Gd\mathbf{r}\cdot\nabla F = d\mathbf{r}\cdot(F\nabla G + G\nabla F). \end{aligned}$$

Now as these equations hold for any direction $d\mathbf{r}$, the $d\mathbf{r}$ may be canceled by (35), p. 165, and the desired result is obtained.

The use of vector notations for treating assigned practical problems involving computation is not great, but for handling the general theory of such parts of physics as are essentially concerned with direct quantities, mechanics, hydromechanics, electromagnetic theories, etc., the actual use of the vector algorisms considerably shortens the formulas and has the added advantage of operating directly upon the magnitudes involved. At this point some of the elements of mechanics will be developed.

79. According to Newton's Second Law, when a force acts upon a particle of mass m, *the rate of change of momentum is equal to the force acting, and takes place in the direction of the force.* It therefore appears that the rate of change of momentum and momentum itself are to be regarded as vector or directed magnitudes in the application of the Second Law. Now if the vector $\mathbf{r}$, laid off from a fixed origin to the point at which the moving mass m is situated at any instant of time t, be differentiated with respect to the time t, the derivative $d\mathbf{r}/dt$ is a vector, tangent to the curve in which the particle is moving and of magnitude equal to ds/dt or v, the velocity of motion. As vectors*, then, the velocity $\mathbf{v}$ and the momentum and the force may be written as

$$\mathbf{v} = \frac{d\mathbf{r}}{dt}, \qquad m\mathbf{v}, \qquad \mathbf{F} = \frac{d}{dt}(m\mathbf{v}). \tag{51}$$

Hence $$\mathbf{F} = m\frac{d\mathbf{v}}{dt} = m\frac{d^2\mathbf{r}}{dt^2} = m\mathbf{f} \quad \text{if} \quad \mathbf{f} = \frac{d\mathbf{v}}{dt} = \frac{d^2\mathbf{r}}{dt^2}.$$

From the equations it appears that the force $\mathbf{F}$ is the product of the mass m by a vector $\mathbf{f}$ which is the rate of change of the velocity regarded

* In applications, it is usual to denote vectors by heavy type and to denote the magnitudes of those vectors by corresponding italic letters.

as a vector. The vector $\mathbf{f}$ is called the *acceleration;* it must not be confused with the rate of change dv/dt or d^2s/dt^2 of the speed or magnitude of the velocity. The components f_x, f_y, f_z of the acceleration along the axes are the projections of $\mathbf{f}$ along the directions $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$ and may be written as $\mathbf{f}\cdot\mathbf{i}$, $\mathbf{f}\cdot\mathbf{j}$, $\mathbf{f}\cdot\mathbf{k}$. Then by the laws of differentiation it follows that

$$f_x = \mathbf{f}\cdot\mathbf{i} = \frac{d\mathbf{v}}{dt}\cdot\mathbf{i} = \frac{d\,(\mathbf{v}\cdot\mathbf{i})}{dt} = \frac{dv_x}{dt},$$

or

$$f_x = \mathbf{f}\cdot\mathbf{i} = \frac{d^2\mathbf{r}}{dt^2}\cdot\mathbf{i} = \frac{d^2\,(\mathbf{r}\cdot\mathbf{i})}{dt^2} = \frac{d^2x}{dt^2}.$$

Hence

$$f_x = \frac{d^2x}{dt^2}, \qquad f_y = \frac{d^2y}{dt^2}, \qquad f_z = \frac{d^2z}{dt^2},$$

and it is seen that the components of the acceleration are the accelerations of the components. If X, Y, Z are the components of the force, the equations of motion in rectangular coördinates are

$$m\frac{d^2x}{dt^2} = X, \qquad m\frac{d^2y}{dt^2} = Y, \qquad m\frac{d^2z}{dt^2} = Z. \tag{52}$$

Instead of resolving the acceleration, force, and displacement along the axes, it may be convenient to resolve them along the tangent and normal to the curve. The velocity $\mathbf{v}$ may be written as $v\mathbf{t}$, where v is the magnitude of the velocity and $\mathbf{t}$ is a unit vector tangent to the curve. Then

$$\mathbf{f} = \frac{d\mathbf{v}}{dt} = \frac{d(v\mathbf{t})}{dt} = \frac{dv}{dt}\mathbf{t} + v\frac{d\mathbf{t}}{dt}.$$

But

$$\frac{d\mathbf{t}}{dt} = \frac{d\mathbf{t}}{ds}\frac{ds}{dt} = \mathbf{C}v = \frac{v}{R}\mathbf{n}, \tag{53}$$

where R is the radius of curvature and $\mathbf{n}$ is a unit normal. Hence

$$\mathbf{f} = \frac{d^2s}{dt^2}\mathbf{t} + \frac{v^2}{R}\mathbf{n}, \qquad f_t = \frac{d^2s}{dt^2}, \qquad f_n = \frac{v^2}{R}. \tag{53'}$$

It therefore is seen that the component of the acceleration along the tangent is d^2s/dt^2, or the rate of change of the velocity regarded as a number, and the component normal to the curve is v^2/R. If T and N are the components of the force along the tangent and normal to the curve of motion, the equations are

$$T = mf_t = m\frac{d^2s}{dt^2}, \qquad N = mf_n = m\frac{v^2}{R}.$$

It is noteworthy that the force must lie in the osculating plane.

If $\mathbf{r}$ and $\mathbf{r} + \Delta\mathbf{r}$ are two positions of the radius vector, the area of the sector included by them is (except for infinitesimals of higher order)

$\Delta\mathbf{A} = \frac{1}{2}\,\mathbf{r}\times(\mathbf{r} + \Delta\mathbf{r}) = \frac{1}{2}\,\mathbf{r}\times\Delta\mathbf{r}$, and is a vector quantity of which the direction is normal to the plane of $\mathbf{r}$ and $\mathbf{r} + \Delta\mathbf{r}$, that is, to the plane through the origin tangent to the curve. The rate of description of area, or the *areal velocity*, is therefore

$$\frac{d\mathbf{A}}{dt} = \lim \tfrac{1}{2}\,\mathbf{r}\times\frac{\Delta\mathbf{r}}{\Delta t} = \tfrac{1}{2}\,\mathbf{r}\times\frac{d\mathbf{r}}{dt} = \tfrac{1}{2}\,\mathbf{r}\times\mathbf{v}. \tag{54}$$

The projections of the areal velocities on the coördinate planes, which are the same as the areal velocities of the projection of the motion on those planes, are (Ex. 11 below)

$$\frac{1}{2}\left(y\frac{dz}{dt} - z\frac{dy}{dt}\right), \qquad \frac{1}{2}\left(z\frac{dx}{dt} - x\frac{dz}{dt}\right), \qquad \frac{1}{2}\left(x\frac{dy}{dt} - y\frac{dx}{dt}\right). \tag{54'}$$

If the force $\mathbf{F}$ acting on the mass m passes through the origin, then $\mathbf{r}$ and $\mathbf{F}$ lie along the same direction and $\mathbf{r}\times\mathbf{F} = 0$. The equation of motion may then be integrated at sight.

$$m\frac{d\mathbf{v}}{dt} = \mathbf{F}, \qquad m\mathbf{r}\times\frac{d\mathbf{v}}{dt} = \mathbf{r}\times\mathbf{F} = 0,$$

$$\mathbf{r}\times\frac{d\mathbf{v}}{dt} = \frac{d}{dt}(\mathbf{r}\times\mathbf{v}) = 0, \qquad \mathbf{r}\times\mathbf{v} = \text{const.}$$

It is seen that in this case the rate of description of area is a constant vector, which means that the rate is not only constant in magnitude but is constant in direction, that is, the path of the particle m must lie in a plane through the origin. When the force passes through a fixed point, as in this case, the force is said to be *central*. Therefore when a particle moves under the action of a central force, the motion takes place in a plane passing through the center and the rate of description of areas, or the areal velocity, is constant.

80. If there are several particles, say n, in motion, each has its own equation of motion. These equations may be combined by addition and subsequent reduction.

$$m_1\frac{d^2\mathbf{r}_1}{dt^2} = \mathbf{F}_1,\; m_2\frac{d^2\mathbf{r}_2}{dt^2} = \mathbf{F}_2,\; \cdots,\; m_n\frac{d^2\mathbf{r}_n}{dt^2} = \mathbf{F}_n,$$

and
$$m_1\frac{d^2\mathbf{r}_1}{dt^2} + m_2\frac{d^2\mathbf{r}_2}{dt^2} + \cdots + m_n\frac{d^2\mathbf{r}_n}{dt^2} = \mathbf{F}_1 + \mathbf{F}_2 + \cdots + \mathbf{F}_n.$$

But
$$m_1\frac{d^2\mathbf{r}_1}{dt^2} + m_2\frac{d^2\mathbf{r}_2}{dt^2} + \cdots + m_n\frac{d^2\mathbf{r}_n}{dt^2} = \frac{d^2}{dt^2}(m_1\mathbf{r}_1 + m_2\mathbf{r}_2 + \cdots + m_n\mathbf{r}_n).$$

Let
$$m_1\mathbf{r}_1 + m_2\mathbf{r}_2 + \cdots + m_n\mathbf{r}_n = (m_1 + m_2 + \cdots + m_n)\,\bar{\mathbf{r}} = M\,\bar{\mathbf{r}}$$

or
$$\bar{\mathbf{r}} = \frac{m_1\mathbf{r}_1 + m_2\mathbf{r}_2 + \cdots + m_n\mathbf{r}_n}{m_1 + m_2 + \cdots + m_n} = \frac{\Sigma m\mathbf{r}}{\Sigma m} = \frac{\Sigma m\mathbf{r}}{M}.$$

Then
$$M\frac{d^2\bar{\mathbf{r}}}{dt^2} = \mathbf{F}_1 + \mathbf{F}_2 + \cdots + \mathbf{F}_n = \sum\mathbf{F}. \tag{55}$$

Now the vector $\mathbf{r}$ which has been here introduced is the vector of the center of mass or center of gravity of the particles (Ex. 5, p. 168). The result (55) states, on comparison with (51), that the center of gravity of the n masses moves as if all the mass M were concentrated at it and all the forces applied there.

The force $\mathbf{F}_i$ acting on the ith mass may be wholly or partly due to attractions, repulsions, pressures, or other actions exerted on that mass by one or more of the other masses of the system of n particles. In fact let $\mathbf{F}_i$ be written as

$$\mathbf{F}_i = \mathbf{F}_{i0} + \mathbf{F}_{i1} + \mathbf{F}_{i2} + \cdots + \mathbf{F}_{in},$$

where $\mathbf{F}_{ij}$ is the force exerted on m_i by m_j and $\mathbf{F}_{i0}$ is the force due to some agency external to the n masses which form the system. Now by Newton's Third Law, when one particle acts upon a second, the second reacts upon the first with a force which is equal in magnitude and opposite in direction. Hence to $\mathbf{F}_{ij}$ above there will correspond a force $\mathbf{F}_{ji} = -\mathbf{F}_{ij}$ exerted by m_i on m_j. In the sum $\Sigma\mathbf{F}_i$ all these equal and opposite actions and reactions will drop out and $\Sigma\mathbf{F}_i$ may be replaced by $\Sigma\mathbf{F}_{i0}$, the sum of the external forces. Hence the important theorem that: *The motion of the center of mass of a set of particles is as if all the mass were concentrated there and all the external forces were applied there* (the internal forces, that is, the forces of mutual action and reaction between the particles being entirely neglected).

The *moment of a force* about a given point is defined as the product of the force by the perpendicular distance of the force from the point. If $\mathbf{r}$ is the vector from the point as origin to any point in the line of the force, the moment is therefore $\mathbf{r}\times\mathbf{F}$ when considered as a vector quantity, and is perpendicular to the plane of the line of the force and the origin. The equations of n moving masses may now be combined in a different way and reduced. Multiply the equations by $\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_n$ and add. Then

$$m_1\mathbf{r}_1\times\frac{d\mathbf{v}_1}{dt} + m_2\mathbf{r}_2\times\frac{d\mathbf{v}_2}{dt} + \cdots + m_n\mathbf{r}_n\times\frac{d\mathbf{v}_n}{dt} = \mathbf{r}_1\times\mathbf{F}_1 + \mathbf{r}_2\times\mathbf{F}_2 + \cdots + \mathbf{r}_n\times\mathbf{F}_n$$

or $$m_1\frac{d}{dt}\mathbf{r}_1\times\mathbf{v}_1 + m_2\frac{d}{dt}\mathbf{r}_2\times\mathbf{v}_2 + \cdots + m_n\frac{d}{dt}\mathbf{r}_n\times\mathbf{v}_n = \mathbf{r}_1\times\mathbf{F}_1 + \mathbf{r}_2\times\mathbf{F}_2 + \cdots + \mathbf{r}_n\times\mathbf{F}_n$$

or $$\frac{d}{dt}(m_1\mathbf{r}_1\times\mathbf{v}_1 + m_2\mathbf{r}_2\times\mathbf{v}_2 + \cdots + m_n\mathbf{r}_n\times\mathbf{v}_n) = \Sigma\mathbf{r}\times\mathbf{F}. \qquad (56)$$

This equation shows that if the areal velocities of the different masses are multiplied by those masses, and all added together, the derivative of the sum obtained is equal to the moment of all the forces about the origin, the moments of the different forces being added as vector quantities.

This result may be simplified and put in a different form. Consider again the resolution of $\mathbf{F}_i$ into the sum $\mathbf{F}_{i0} + \mathbf{F}_{i1} + \cdots + \mathbf{F}_{in}$, and in particular consider the action $\mathbf{F}_{ij}$ and the reaction $\mathbf{F}_{ji} = -\mathbf{F}_{ij}$ between two particles. Let it be assumed that the action and reaction are not only equal and opposite, but lie along the line connecting the two particles. Then the perpendicular distances from the origin to the action and reaction are equal and the moments of the action and reaction are equal and opposite, and may be dropped from the sum $\Sigma\mathbf{r}_i\times\mathbf{F}_i$, which then reduces to $\Sigma\mathbf{r}_i\times\mathbf{F}_{i0}$. On the other hand a term like $m_i\mathbf{r}_i\times\mathbf{v}_i$ may be written as $\mathbf{r}_i\times(m_i\mathbf{v}_i)$. This product is formed from the momentum in exactly the same way that the moment is formed from the force, and it is called the moment of momentum. Hence the equation (56) becomes

$$\frac{d}{dt}\text{(total moment of momentum)} = \text{moment of external forces.}$$

Hence the result that, as vector quantities: *The rate of change of the moment of momentum of a system of particles is equal to the moment of the external forces* (the forces between the masses being entirely neglected under the assumption that action and reaction lie along the line connecting the masses).

EXERCISES

1. Apply the definition of differentiation to prove

(α) $d(\mathbf{u}\cdot\mathbf{v}) = \mathbf{u}\cdot d\mathbf{v} + \mathbf{v}\cdot d\mathbf{u}$, (β) $d[\mathbf{u}\cdot(\mathbf{v}\times\mathbf{w})] = d\mathbf{u}\cdot(\mathbf{v}\times\mathbf{w}) + \mathbf{u}\cdot(d\mathbf{v}\times\mathbf{w}) + \mathbf{u}\cdot(\mathbf{v}\times d\mathbf{w})$.

2. Differentiate under the assumption that vectors denoted by early letters of the alphabet are constant and those designated by the later letters are variable:

(α) $\mathbf{u}\times(\mathbf{v}\times\mathbf{w})$, (β) $\mathbf{a}\cos t + \mathbf{b}\sin t$, (γ) $(\mathbf{u}\cdot\mathbf{u})\,\mathbf{u}$,

(δ) $\mathbf{u}\times\dfrac{d\mathbf{u}}{dx}$, (ϵ) $\mathbf{u}\cdot\left(\dfrac{d\mathbf{u}}{dx}\times\dfrac{d^2\mathbf{u}}{dx^2}\right)$, (ζ) $\mathbf{c}(\mathbf{a}\cdot\mathbf{u})$.

3. Apply the rules for change of variable to show that $\dfrac{d^2\mathbf{r}}{ds^2} = \dfrac{\mathbf{r}''s' - \mathbf{r}'s''}{s'^3}$, where accents denote differentiation with respect to x. In case $\mathbf{r} = x\mathbf{i} + y\mathbf{j}$ show that $1/\sqrt{\mathbf{C}\cdot\mathbf{C}}$ takes the usual form for the radius of curvature of a plane curve.

4. The equation of the helix is $r = \mathbf{i}a\cos\phi + \mathbf{j}a\sin\phi + \mathbf{k}b\phi$ with $s = \sqrt{a^2+b^2}\,\phi$; show that the radius of curvature is $(a^2+k^2)/a$.

5. Find the torsion of the helix. It is $b/(a^2+k^2)$.

6. Change the variable from s to some other variable t in the formula for torsion.

7. In the following cases find the gradient either by applying the formula which contains the partial derivatives, or by using the relation $d\mathbf{r}\cdot\nabla F = dF$, or both:

(α) $\mathbf{r}\cdot\mathbf{r} = x^2+y^2+z^2$, (β) $\log r$, (γ) $r = \sqrt{\mathbf{r}\cdot\mathbf{r}}$,

(δ) $\log(x^2+y^2) = \log[\mathbf{r}\cdot\mathbf{r} - (\mathbf{k}\cdot\mathbf{r})^2]$, (ϵ) $(\mathbf{r}\times\mathbf{a})\cdot(\mathbf{r}\times\mathbf{b})$.

8. Prove these laws of operation with the symbol ∇:

(α) $\nabla(F+G) = \nabla F + \nabla G$, (β) $G^2\nabla(F/G) = G\nabla F - F\nabla G$.

9. If r, ϕ are polar coördinates in a plane and $\mathbf{r}_1$ is a unit vector along the radius vector, show that $d\mathbf{r}_1/dt = \mathbf{n}d\phi/dt$ where $\mathbf{n}$ is a unit vector perpendicular to the radius. Thus differentiate $\mathbf{r} = r\mathbf{r}_1$ twice and separate the result into components along the radius vector and perpendicular to it so that

$$f_r = \frac{d^2r}{dt^2} - r\left(\frac{d\phi}{dt}\right)^2, \qquad f_\phi = r\frac{d^2\phi}{dt^2} + 2\frac{d\phi}{dt}\frac{dr}{dt} = \frac{1}{r}\frac{d}{dt}\left(r^2\frac{d\phi}{dt}\right).$$

10. Prove conversely to the text that if the vector rate of description of area is constant, the force must be central, that is, $\mathbf{r}\times\mathbf{F} = 0$.

11. Note that $\mathbf{r}\times\mathbf{v}\cdot\mathbf{i}$, $\mathbf{r}\times\mathbf{v}\cdot\mathbf{j}$, $\mathbf{r}\times\mathbf{v}\cdot\mathbf{k}$ are the projections of the areal velocities upon the planes $x=0$, $y=0$, $z=0$. Hence derive (54′) of the text.

12. Show that the Cartesian expressions for the magnitude of the velocity and of the acceleration and for the rate of change of the speed dv/dt are

$$v = \sqrt{x'^2 + y'^2 + z'^2}, \quad f = \sqrt{x''^2 + y''^2 + z''^2}, \quad v' = \frac{x'x'' + y'y'' + z'z''}{\sqrt{x'^2 + y'^2 + z'^2}},$$

where accents denote differentiation with respect to the time.

13. Suppose that a body which is rigid is rotating about an axis with the angular velocity $\omega = d\phi/dt$. Represent the angular velocity by a vector $\mathbf{a}$ drawn along the axis and of magnitude equal to ω. Show that the velocity of any point in space is $\mathbf{v} = \mathbf{a}\times\mathbf{r}$, where $\mathbf{r}$ is the vector drawn to that point from any point of the axis as origin. Show that the acceleration of the point determined by $\mathbf{r}$ is in a plane through the point and perpendicular to the axis, and that the components are

$$\mathbf{a}\times(\mathbf{a}\times\mathbf{r}) = (\mathbf{a}\cdot\mathbf{r})\,\mathbf{a} - \omega^2\mathbf{r} \text{ toward the axis}, \quad (d\mathbf{a}/dt)\times\mathbf{r} \text{ perpendicular to the axis},$$

under the assumption that the axis of rotation is invariable.

14. Let $\bar{\mathbf{r}}$ denote the center of gravity of a system of particles and $\mathbf{r}_i'$ denote the vector drawn from the center of gravity to the ith particle so that $\mathbf{r}_i = \bar{\mathbf{r}} + \mathbf{r}_i'$ and $\mathbf{v}_i = \bar{\mathbf{v}} + \mathbf{v}_i'$. The kinetic energy of the ith particle is by definition

$$\tfrac{1}{2}m_iv_i^2 = \tfrac{1}{2}m_i\mathbf{v}_i\cdot\mathbf{v}_i = \tfrac{1}{2}m_i(\bar{\mathbf{v}} + \mathbf{v}_i')\cdot(\bar{\mathbf{v}} + \mathbf{v}_i').$$

Sum up for all particles and simplify by using the fact $\Sigma m_i\mathbf{r}_i' = 0$, which is due to the assumption that the origin for the vectors $\mathbf{r}_i'$ is at the center of gravity. Hence prove the important theorem: *The total kinetic energy of a system is equal to the kinetic energy which the total mass would have if moving with the center of gravity plus the energy computed from the motion relative to the center of gravity as origin*, that is,

$$T = \tfrac{1}{2}\Sigma m_iv_i^2 = \tfrac{1}{2}M\bar{v}^2 + \tfrac{1}{2}\Sigma m_iv_i'^2.$$

15. Consider a rigid body moving in a plane, which may be taken as the xy-plane. Let any point $\mathbf{r}_0$ of the body be marked and other points be denoted relative to it by $\mathbf{r}'$. The motion of any point $\mathbf{r}'$ is compounded from the motion of $\mathbf{r}_0$ and from the angular velocity $\mathbf{a} = \mathbf{k}\omega$ of the body about the point $\mathbf{r}_0$. In fact the velocity $\mathbf{v}$ of any point is $\mathbf{v} = \mathbf{v}_0 + \mathbf{a}\times\mathbf{r}$. Show that the velocity of the point denoted by $\mathbf{r}' = \mathbf{k}\times\mathbf{v}_0/\omega$ is zero. This point is known as the instantaneous center of rotation (§ 39). Show that the coördinates of the instantaneous center referred to axes at the origin of the vectors $\mathbf{r}$ are

$$x = \mathbf{r}\cdot\mathbf{i} = x_0 - \frac{1}{\omega}\frac{dy_0}{dt}, \qquad y = \mathbf{r}\cdot\mathbf{j} = y_0 + \frac{1}{\omega}\frac{dx_0}{dt}.$$

16. If several forces $\mathbf{F}_1, \mathbf{F}_2, \cdots, \mathbf{F}_n$ act on a body, the sum $\mathbf{R} = \Sigma\mathbf{F}_i$ is called the *resultant* and the sum $\Sigma\mathbf{r}_i\times\mathbf{F}_i$, where $\mathbf{r}_i$ is drawn from an origin O to a point in the line of the force $\mathbf{F}_i$, is called the *resultant moment* about O. Show that the resultant moments $\mathbf{M}_O$ and $\mathbf{M}_{O'}$ about two points are connected by the relation $\mathbf{M}_{O'} = \mathbf{M}_O + \mathbf{M}_{O'}(\mathbf{R}_O)$, where $\mathbf{M}_{O'}(\mathbf{R}_O)$ means the moment about O' of the resultant $\mathbf{R}$ considered as applied at O. Infer that moments about all points of any line parallel to the resultant are equal. Show that in any plane perpendicular to $\mathbf{R}$ there is a point O' given by $\mathbf{r} = \mathbf{R}\times\mathbf{M}_O/\mathbf{R}\cdot\mathbf{R}$, where O is any point of the plane, such that $\mathbf{M}_{O'}$ is parallel to $\mathbf{R}$.

PART II. DIFFERENTIAL EQUATIONS

CHAPTER VII

GENERAL INTRODUCTION TO DIFFERENTIAL EQUATIONS

81. Some geometric problems. The application of the differential calculus to plane curves has given a means of determining some geometric properties of the curves. For instance, the length of the subnormal of a curve (§ 7) is ydy/dx, which in the case of the parabola $y^2 = 4px$ is $2p$, that is, the subnormal is constant. Suppose now it were desired conversely to find all curves for which the subnormal is a given constant m. The statement of this problem is evidently contained in the equation

$$y\frac{dy}{dx} = m \quad \text{or} \quad yy' = m \quad \text{or} \quad ydy = mdx.$$

Again, the radius of curvature of the lemniscate $r^2 = a^2 \cos 2\phi$ is found to be $R = a^2/3r$, that is, the radius of curvature varies inversely as the radius. If conversely it were desired to find all curves for which the radius of curvature varies inversely as the radius of the curve, the statement of the problem would be the equation

$$\frac{\left[r^2 + \left(\frac{dr}{d\phi}\right)^2\right]^{\frac{3}{2}}}{r^2 - r\frac{d^2r}{d\phi^2} + 2\left(\frac{dr}{d\phi}\right)^2} = \frac{k}{r},$$

where k is a constant called a factor of proportionality.*

Equations like these are unlike ordinary algebraic equations because, in addition to the variables x, y or r, ϕ and certain constants m or k, they contain also derivatives, as dy/dx or $dr/d\phi$ and $d^2r/d\phi^2$, of one of the variables with respect to the other. An equation which contains

* Many problems in geometry, mechanics, and physics are stated in terms of variation. For purposes of analysis the statement x varies as y, or $x \propto y$, is written as $x = ky$, introducing a constant k called a factor of proportionality to convert the variation into an equation. In like manner the statement x varies inversely as y, or $x \propto 1/y$, becomes $x = k/y$, and x varies jointly with y and z becomes $x = kyz$.

derivatives is called a *differential equation.* The *order* of the differential equation is the order of the highest derivative it contains. The equations above are respectively of the first and second orders. A differential equation of the first order may be symbolized as $\Phi(x, y, y') = 0$, and one of the second order as $\Phi(x, y, y', y'') = 0$. A function $y = f(x)$ given explicitly or defined implicitly by the relation $F(x, y) = 0$ is said to be a *solution* of a given differential equation if the equation is true for all values of the independent variable x when the expressions for y and its derivatives are substituted in the equation.

Thus to show that (no matter what the value of a is) the relation

$$4ay - x^2 + 2a^2 \log x = 0$$

gives a solution of the differential equation of the second order

$$1 + \left(\frac{dy}{dx}\right)^2 - x^2\left(\frac{d^2y}{dx^2}\right)^2 = 0,$$

it is merely necessary to form the derivatives

$$2a\frac{dy}{dx} = x - \frac{a^2}{x}, \qquad 2a\frac{d^2y}{dx^2} = 1 + \frac{a^2}{x^2}$$

and substitute them in the given equation together with y to see that

$$1 + \left(\frac{dy}{dx}\right)^2 - x^2\left(\frac{d^2y}{dx^2}\right)^2 = 1 + \frac{1}{4a^2}\left(x^2 - 2a^2 + \frac{a^4}{x^2}\right) - \frac{x^2}{4a^2}\left(1 + \frac{2a^2}{x^2} + \frac{a^4}{x^4}\right) = 0$$

is clearly satisfied for all values of x. It appears therefore that the given relation for y is a solution of the given equation.

To *integrate* or *solve* a differential equation is to find all the functions which satisfy the equation. Geometrically speaking, it is to find all the curves which have the property expressed by the equation. In mechanics it is to find all possible motions arising from the given forces. The method of integrating or solving a differential equation depends largely upon the *ingenuity* of the solver. In many cases, however, some method is immediately obvious. For instance if it be possible to *separate the variables,* so that the differential dy is multiplied by a function of y alone and dx by a function of x alone, as in the equation

$$\phi(y)\,dy = \psi(x)\,dx, \quad \text{then} \quad \int \phi(y)\,dy = \int \psi(x)\,dx + C \qquad (1)$$

will clearly be the integral or solution of the differential equation.

As an example, let the curves of constant subnormal be determined. Here

$$ydy = mdx \quad \text{and} \quad y^2 = 2mx + C.$$

The variables are already separated and the integration is immediate. The curves are parabolas with semi-latus rectum equal to the constant and with the axis coincident with the axis of x. If in particular it were desired to determine that curve whose subnormal was m and which passed through the origin, it would merely be necessary to substitute (0, 0) in the equation $y^2 = 2mx + C$ to ascertain what particular value must be assigned to C in order that the curve pass through (0, 0). The value is $C = 0$.

Another example might be to determine the curves for which the x-intercept varies as the abscissa of the point of tangency. As the expression (§ 7) for the x-intercept is $x - ydx/dy$, the statement is

$$x - y\frac{dx}{dy} = kx \quad \text{or} \quad (1-k)x = y\frac{dx}{dy}.$$

Hence
$$(1-k)\frac{dy}{y} = \frac{dx}{x} \quad \text{and} \quad (1-k)\log y = \log x + C.$$

If desired, this expression may be changed to another form by using each side of the equality as an exponent with the base e. Then

$$e^{(1-k)\log y} = e^{\log x + C} \quad \text{or} \quad y^{1-k} = e^C x = C'x.$$

As C is an arbitrary constant, the constant $C' = e^C$ is also arbitrary and the solution may simply be written as $y^{1-k} = Cx$, where the accent has been omitted from the constant. If it were desired to pick out that particular curve which passed through the point (1, 1), it would merely be necessary to determine C from the equation

$$1^{1-k} = C\,1, \quad \text{and hence} \quad C = 1.$$

As a third example let the curves whose tangent is constant and equal to a be determined. The length of the tangent is $y\sqrt{1+y'^2}/y'$ and hence the equation is

$$y\frac{\sqrt{1+y'^2}}{y'} = a \quad \text{or} \quad y^2\frac{1+y'^2}{y'^2} = a \quad \text{or} \quad 1 = \frac{\sqrt{a^2-y^2}}{y}y'.$$

The variables are therefore separable and the results are

$$dx = \frac{\sqrt{a^2-y^2}}{y}dy \quad \text{and} \quad x + C = \sqrt{a^2-y^2} - a\log\frac{a+\sqrt{a^2-y^2}}{y}.$$

If it be desired that the tangent at the origin be vertical so that the curve passes through $(0, a)$, the constant C is 0. The curve is the tractrix or "curve of pursuit" as described by a calf dragged at the end of a rope by a person walking along a straight line.

82. Problems which involve the radius of curvature will lead to differential equations of the second order. The method of solving such problems is to *reduce the equation, if possible, to one of the first order.* For the second derivative may be written as

$$y'' = \frac{dy'}{dx} = \frac{dy'}{dy}y', \tag{2}$$

and
$$R = \frac{(1+y'^2)^{\frac{3}{2}}}{y''} = \frac{(1+y'^2)^{\frac{3}{2}}}{\frac{dy'}{dx}} = \frac{(1+y'^2)^{\frac{3}{2}}}{y'\frac{dy'}{dy}} \tag{2'}$$

is the expression for the radius of curvature. If it be given that the radius of curvature is of the form $f(x)\phi(y')$ or $f(y)\phi(y')$,

$$\frac{(1+y'^2)^{\frac{3}{2}}}{\frac{dy'}{dx}}=f(x)\phi(y') \quad \text{or} \quad \frac{(1+y'^2)^{\frac{3}{2}}}{y\frac{dy'}{dy}}=f(y)\phi(y'), \tag{3}$$

the variables x and y' or y and y' are immediately separable, and an integration may be performed. This will lead to an equation of the first order; and if the variables are again separable, the solution may be completed by the methods of the above examples.

In the first place consider curves whose radius of curvature is constant. Then

$$\frac{(1+y'^2)^{\frac{3}{2}}}{\frac{dy'}{dx}}=a \quad \text{or} \quad \frac{dy'}{(1+y'^2)^{\frac{3}{2}}}=\frac{dx}{a} \quad \text{and} \quad \frac{y'}{\sqrt{1+y'^2}}=\frac{x-C}{a},$$

where the constant of integration has been written as $-C/a$ for future convenience. The equation may now be solved for y' and the variables become separated with the results

$$y'=\frac{x-C}{\sqrt{a^2-(x-C)^2}} \quad \text{or} \quad dy=\frac{(x-C)}{\sqrt{a^2-(x-C)^2}}dx.$$

Hence $\quad y-C'=-\sqrt{a^2-(x-C)^2} \quad$ or $\quad (x-C)^2+(y-C')^2=a^2.$

The curves, as should be anticipated, are circles of radius a and with any arbitrary point (C, C') as center. It should be noted that, as the solution has required two successive integrations, there are two arbitrary constants C and C' of integration in the result.

As a second example consider the curves whose radius of curvature is double the normal. As the length of the normal is $y\sqrt{1+y'^2}$, the equation becomes

$$\frac{(1+y'^2)^{\frac{3}{2}}}{y'\frac{dy'}{dy}}=2y\sqrt{1+y'^2} \quad \text{or} \quad \frac{1+y'^2}{y'\frac{dy'}{dy}}=\pm 2y,$$

where the double sign has been introduced when the radical is removed by cancellation. This is necessary; for before the cancellation the signs were ambiguous and there is no reason to assume that the ambiguity disappears. In fact, if the curve is concave up, the second derivative is positive and the radius of curvature is reckoned as positive, whereas the normal is positive or negative according as the curve is above or below the axis of x; similarly, if the curve is concave down. Let the negative sign be chosen. This corresponds to a curve above the axis and concave down, or below the axis and concave up, that is, the normal and the radius of curvature have the same direction. Then

$$\frac{dy}{y}=-\frac{2y'dy'}{1+y'^2} \quad \text{and} \quad \log y=-\log(1+y'^2)+\log 2C,$$

where the constant has been given the form $\log 2C$ for convenience. This expression may be thrown into algebraic form by exponentiation, solved for y', and then

$$y(1+y'^2)=2C \quad \text{or} \quad y'^2=\frac{2C-y}{y} \quad \text{or} \quad \frac{ydy}{\sqrt{2Cy-y^2}}=dx.$$

Hence $$x-C'=C\,\text{vers}^{-1}\frac{y}{C}-\sqrt{2Cy-y^2}.$$

The curves are cycloids of which the generating circle has an arbitrary radius C and of which the cusps are upon the x-axis at the points $C' \pm 2k\pi C$. If the positive sign had been taken in the equation, the curves would have been entirely different; see Ex. 5 (α).

The number of arbitrary constants of integration which enter into the solution of a differential equation depends on the number of integrations which are performed and is equal to the order of the equation. This results in giving a family of curves, dependent on one or more parameters, as the solution of the equation. To pick out any particular member of the family, additional conditions must be given. Thus, if there is only one constant of integration, the curve may be required to pass through a given point; if there are two constants, the curve may be required to pass through a given point and have a given slope at that point, or to pass through two given points. These additional conditions are called *initial conditions*. In mechanics the initial conditions are very important; for the point reached by a particle describing a curve under the action of assigned forces depends not only on the forces, but on the point at which the particle started and the velocity with which it started. In all cases the distinction between the *constants of integration* and the *given constants of the problem* (in the foregoing examples, the distinction between C, C' and m, k, a) should be kept clearly in mind

EXERCISES

1. Verify the solutions of the differential equations:

(α) $xy+\frac{1}{2}x^2=C,\ y+x+xy'=0,$ (β) $x^3y^3(3e^x+C)=1,\ xy'+y+x^4y^4e^x=0,$

(γ) $(1+x^2)y'^2=1,\ 2x=Ce^y-C^{-1}e^{-y},$ (δ) $y+xy'=x^4y'^2,\ xy=C^2x+C,$

(ϵ) $y''+y'/x=0,\ y=C\log x+C_1,$ (ζ) $y=Ce^x+C_1e^{2x},\ y''+2y=3y',$

(η) $y'''-y=x^2,\ y=Ce^x+e^{-\frac{1}{2}x}\left(C_1\cos\frac{x\sqrt{3}}{2}+C_2\sin\frac{x\sqrt{3}}{2}\right)-x^2.$

2. Determine the curves which have the following properties:

(α) The subtangent is constant; $y=Cx^m$. If through (2, 2), $y=2^{1-m}x^m$.

(β) The right triangle formed by the tangent, subtangent, and ordinate has the constant area $k/2$; the hyperbolas $xy+Cy+k=0$. Show that if the curve passes through (1, 2) and (2, 1), the arbitrary constant C is 0 and the given k is -2.

(γ) The normal is constant in length; the circles $(x-C)^2+y^2=k^2$.

(δ) The normal varies as the square of the ordinate; catenaries $ky=\cosh k(x-C)$. If in particular the curve is perpendicular to the y-axis, $C=0$.

(ϵ) The area of the right triangle formed by the tangent, normal, and x-axis is inversely proportional to the slope; the circles $(x-C)^2+y^2=k$.

3. Determine the curves which have the following properties:

(α) The angle between the radius vector and tangent is constant; spirals $r = Ce^{k\phi}$.

(β) The angle between the radius vector and tangent is half that between the radius and initial line; cardioids $r = C(1 - \cos\phi)$.

(γ) The perpendicular from the pole to a tangent is constant; $r\cos(\phi - C) = k$.

(δ) The tangent is equally inclined to the radius vector and to the initial line; the two sets of parabolas $r = C/(1 \pm \cos\phi)$.

(ϵ) The radius is equally inclined to the normal and to the initial line; circles $r = C\cos\phi$ or lines $r\cos\phi = C$.

4. The arc s of a curve is proportional to the area A, where in rectangular coördinates A is the area under the curve and in polar coördinates it is the area included by the curve and the radius vectors. From the equation $ds = dA$ show that the curves which satisfy the condition are catenaries for rectangular coördinates and lines for polar coördinates.

5. Determine the curves for which the radius of curvature

(α) is twice the normal and oppositely directed; parabolas $(x - C)^2 = C'(2y - C')$.

(β) is equal to the normal and in same direction; circles $(x - C)^2 + y^2 = C'^2$.

(γ) is equal to the normal and in opposite direction; catenaries.

(δ) varies as the cube of the normal; conics $kCy^2 - C^2(x + C')^2 = k$.

(ϵ) projected on the x-axis equals the abscissa; circles.

(ζ) projected on the x-axis is the negative of the abscissa; catenaries.

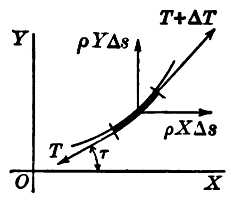
(η) projected on the x-axis is twice the abscissa; central conics.

(θ) is proportional to the slope of the tangent or of the normal.

83. Problems in mechanics and physics. In many physical problems the statement involves an equation between the *rate of change* of some quantity and the value of that quantity. In this way the solution of the problem is made to depend on the integration of a differential equation of the first order. If x denotes any quantity, the rate of increase in x is dx/dt and the rate of decrease in x is $-dx/dt$; and consequently when the rate of change of x is a function of x, the variables are immediately separated and the integration may be performed. The constant of integration has to be determined from the initial conditions; the constants inherent in the problem may be given in advance or their values may be determined by comparing x and t at some subsequent time. The exercises offered below will exemplify the treatment of such problems.

In other physical problems the statement of the question as a differential equation is not so direct and is carried out by an examination of the problem with a view to stating a relation between the increments or differentials of the dependent and independent variables, as in some geometric relations already discussed (§ 40), and in the problem of the tension in a rope wrapped around a cylindrical post discussed below.

The method may be further illustrated by the derivation of the differential equations of the curve of equilibrium of a flexible string or chain. Let ρ be the density of the chain so that $\rho\Delta s$ is the mass of the length Δs; let X and Y be the components of the force (estimated per unit mass) acting on the elements of the chain. Let T denote the tension in the chain, and τ the inclination of the element of chain. From the figure it then appears that the components of all the forces acting on Δs are



$$(T + \Delta T) \cos (\tau + \Delta\tau) - T \cos \tau + X\rho\Delta s = 0,$$
$$(T + \Delta T) \sin (\tau + \Delta\tau) - T \sin \tau + Y\rho\Delta s = 0;$$

for these must be zero if the element is to be in a position of equilibrium. The equations may be written in the form

$$\Delta (T \cos \tau) + X\rho\Delta s = 0, \qquad \Delta (T \sin \tau) + Y\rho\Delta s = 0;$$

and if they now be divided by Δs and if Δs be allowed to approach zero, the result is the two equations of equilibrium

$$\frac{d}{ds}\left(T\frac{dx}{ds}\right) + \rho X = 0, \qquad \frac{d}{ds}\left(T\frac{dy}{ds}\right) + \rho Y = 0, \tag{4}$$

where $\cos \tau$ and $\sin \tau$ are replaced by their values dx/ds and dy/ds.

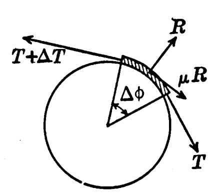
If the string is acted on only by forces parallel to a given direction, let the y-axis be taken as parallel to that direction. Then the component X will be zero and the first equation may be integrated. The result is

$$\frac{d}{ds}\left(T\frac{dx}{ds}\right) = 0, \qquad T\frac{dx}{ds} = C, \qquad T = C\frac{ds}{dx}.$$

This value of T may be substituted in the second equation. There is thus obtained a differential equation of the second order

$$\frac{d}{ds}\left(C\frac{dy}{dx}\right) + \rho Y = 0 \quad \text{or} \quad C\frac{y''}{\sqrt{1 + y'^2}} + \rho Y = 0. \tag{4'}$$

If this equation can be integrated, the form of the curve of equilibrium may be found.



Another problem of a different nature in strings is to consider the variation of the tension in a rope wound around a cylinder without overlapping. The forces acting on the element Δs of the rope are the tensions T and $T + \Delta T$, the normal pressure or reaction R of the cylinder, and the force of friction which is proportional to the pressure. It will be assumed that the normal reaction lies in the angle $\Delta\phi$ and that the coefficient of friction is μ so that the force of friction is μR. The components along the radius and along the tangent are

$$(T + \Delta T) \sin \Delta\phi - R \cos(\theta\Delta\phi) - \mu R \sin(\theta\Delta\phi) = 0, \qquad 0 < \theta < 1,$$
$$(T + \Delta T) \cos \Delta\phi + R \sin(\theta\Delta\phi) - \mu R \cos(\theta\Delta\phi) - T = 0.$$

Now discard all infinitesimals except those of the first order. It must be borne in mind that the pressure R is the reaction on the infinitesimal arc Δs and hence is itself infinitesimal. The substitutions are therefore $Td\phi$ for $(T + \Delta T) \sin \Delta\phi$, R for $R \cos \theta\Delta\phi$, 0 for $R \sin \theta\Delta\phi$, and $T + dT$ for $(T + \Delta T) \cos \Delta\phi$. The equations therefore reduce to two simple equations

$$Td\phi - R = 0, \qquad dT - \mu R = 0,$$

from which the unknown R may be eliminated with the result

$$dT = \mu T d\phi \quad \text{or} \quad T = Ce^{\mu\phi} \quad \text{or} \quad T = T_0 e^{\mu\phi},$$

where T_0 is the tension when ϕ is 0. The tension therefore runs up exponentially and affords ample explanation of why a man, by winding a rope about a post, can readily hold a ship or other object exerting a great force at the other end of the rope. If μ is 1/3, three turns about the post will hold a force 535 T_0, or over 25 tons, if the man exerts a force of a hundredweight.

84. If a constant mass m is moving along a line under the influence of a force F acting along the line, Newton's Second Law of Motion (p. 13) states the problem of the motion as the differential equation

$$mf = F \quad \text{or} \quad m\frac{d^2x}{dt^2} = F \tag{5}$$

of the second order; and it therefore appears that the complete solution of a problem in rectilinear motion requires the integration of this equation. The acceleration may be written as

$$f = \frac{dv}{dt} = \frac{dv}{dx}\frac{dx}{dt} = v\frac{dv}{dx};$$

and hence the equation of motion takes either of the forms

$$m\frac{dv}{dt} = F \quad \text{or} \quad mv\frac{dv}{dx} = F. \tag{5'}$$

It now appears that there are several cases in which the first integration may be performed. For if the force is a function of the velocity or of the time or a product of two such functions, the variables are separated in the first form of the equation; whereas if the force is a function of the velocity or of the coördinate x or a product of two such functions, the variables are separated in the second form of the equation.

When the first integration is performed according to either of these methods, there will arise an equation between the velocity and either the time t or the coördinate x. In this equation will be contained a constant of integration which may be determined by the initial conditions, that is, by the knowledge of the velocity at the start, whether in time or in position. Finally it will be possible (at least theoretically) to solve the equation and express the velocity as a function of the time t or of the position x, as the case may be, and integrate a second time. The carrying through in practice of this sketch of the work will be exemplified in the following two examples.

Suppose a particle of mass m is projected vertically upward with the velocity V. Solve the problem of the motion under the assumption that the resistance of the air varies as the velocity of the particle. Let the distance be measured vertically upward. The forces acting on the particle are two, — the force of gravity which is the weight $W = mg$, and the resistance of the air which is kv. Both these forces are negative because they are directed toward diminishing values of x. Hence

$$mf = -mg - kv \quad \text{or} \quad m\frac{dv}{dt} = -mg - kv,$$

where the first form of the equation of motion has been chosen, although in this case the second form would be equally available. Then integrate.

$$\frac{dv}{g + \frac{k}{m}v} = -dt \quad \text{and} \quad \log\left(g + \frac{k}{m}v\right) = -\frac{k}{m}t + C.$$

As by the initial conditions $v = V$ when $t = 0$, the constant C is found from

$$\log\left(g + \frac{k}{m}V\right) = -\frac{k}{m}0 + C; \quad \text{hence} \quad \frac{g + \frac{k}{m}v}{g + \frac{k}{m}V} = e^{-\frac{k}{m}t}$$

is the relation between v and t found by substituting the value of C. The solution for v gives

$$v = \frac{dx}{dt} = \left(\frac{m}{k}g + V\right)e^{-\frac{k}{m}t} - \frac{m}{k}g.$$

Hence $$x = -\frac{m}{k}\left(\frac{m}{k}g + V\right)e^{-\frac{k}{m}t} - \frac{m}{k}gt + C.$$

If the particle starts from the origin $x = 0$, the constant C is found to be

$$C = \frac{m}{k}\left(\frac{m}{k}g + V\right) \quad \text{and} \quad x = \frac{m}{k}\left(\frac{m}{k}g + V\right)\left(1 - e^{-\frac{k}{m}t}\right) - \frac{m}{k}gt.$$

Hence the position of the particle is expressed in terms of the time and the problem is solved. If it be desired to find the time which elapses before the particle comes to rest and starts to drop back, it is merely necessary to substitute $v = 0$ in the relation connecting the velocity and the time, and solve for the time $t = T$; and if this value of t be substituted in the expression for x, the total distance X covered in the ascent will be found. The results are

$$T = \frac{m}{k}\log\left(1 + \frac{k}{mg}V\right), \qquad X = \left(\frac{m}{k}\right)^2\left[\frac{k}{m}V - g\log\left(1 + \frac{k}{mg}V\right)\right].$$

As a second example consider the motion of a particle vibrating up and down at the end of an elastic string held in the field of gravity. By Hooke's Law for elastic strings the force exerted by the string is proportional to the extension of the string over its natural length, that is, $F = k\Delta l$. Let l be the length of the string, $\Delta_0 l$ the extension of the string just sufficient to hold the weight $W = mg$ at rest so that $k\Delta_0 l = mg$, and let x measured downward be the additional extension of the string at any instant of the motion. The force of gravity mg is positive and the force of elasticity $-k(\Delta_0 l + x)$ is negative. The second form of the equation of motion is to be chosen. Hence

$$mv\frac{dv}{dx} = mg - k(\Delta_0 l + x) \quad \text{or} \quad mv\frac{dv}{dx} = -kx, \quad \text{since} \quad mg = k\Delta_0 l.$$

Then $$mvdv = -kxdx \quad \text{or} \quad mv^2 = -kx^2 + C.$$

Suppose that $x = a$ is the amplitude of the motion, so that when $x = a$ the velocity $v = 0$ and the particle stops and starts back. Then $C = ka^2$. Hence

$$v = \frac{dx}{dt} = \sqrt{\frac{k}{m}}\sqrt{a^2 - x^2} \quad \text{or} \quad \frac{dx}{\sqrt{a^2 - x^2}} = \sqrt{\frac{k}{m}}\,dt,$$

and $$\sin^{-1}\frac{x}{a} = \sqrt{\frac{k}{m}}\,t + C \quad \text{or} \quad x = a\sin\left(\sqrt{\frac{k}{m}}\,t + C\right).$$

Now let the time be measured from the instant when the particle passes through the position $x = 0$. Then C satisfies the equation $0 = a \sin C$ and may be taken as zero. The motion is therefore given by the equation $x = a\sin\sqrt{k/m}\,t$ and is periodic. While t changes by $2\pi\sqrt{m/k}$ the particle completes an entire oscillation. The time $T = 2\pi\sqrt{m/k}$ is called the *periodic time*. The motion considered in this example is characterized by the fact that the total force $-kx$ is proportional to the displacement from a certain origin and is directed toward the origin. Motion of this sort is called *simple harmonic motion* (briefly S. H. M.) and is of great importance in mechanics and physics.

EXERCISES

1. The sum of \$100 is put at interest at 4 per cent per annum under the condition that the interest shall be compounded at each instant. Show that the sum will amount to \$200 in 17 yr. 4 mo., and to \$1000 in 56 yr.

2. Given that the rate of decomposition of an amount x of a given substance is proportional to the amount of the substance remaining undecomposed. Solve the problem of the decomposition and determine the constant of integration and the physical constant of proportionality if $x = 5.11$ when $t = 0$ and $x = 1.48$ when $t = 40$ min. *Ans.* $k = .0309$.

3. A substance is undergoing transformation into another at a rate which is assumed to be proportional to the amount of the substance still remaining untransformed. If that amount is 35.6 when $t = 1$ hr. and 13.8 when $t = 4$ hr., determine the amount at the start when $t = 0$ and the constant of proportionality and find how many hours will elapse before only one-thousandth of the original amount will remain.

4. If the activity A of a radioactive deposit is proportional to its rate of diminution and is found to decrease to $\frac{1}{2}$ its initial value in 4 days, show that A satisfies the equation $A/A_0 = e^{-0.173t}$.

5. Suppose that amounts a and b respectively of two substances are involved in a reaction in which the velocity of transformation dx/dt is proportional to the product $(a-x)(b-x)$ of the amounts remaining untransformed. Integrate on the supposition that $a \neq b$.

$$\log \frac{b(a-x)}{a(b-x)} = (a-b)kt; \quad \text{and if}$$

t	$a-x$	$b-x$
393	0.4866	0.2342
1265	0.3879	0.1354

determine the product $k(a-b)$.

6. Integrate the equation of Ex. 5 if $a = b$, and determine a and k if $x = 9.87$ when $t = 15$ and $x = 13.69$ when $t = 55$.

7. If the velocity of a chemical reaction in which three substances are involved is proportional to the continued product of the amounts of the substances remaining, show that the equation between x and the time is

$$\frac{\log\left(\frac{a}{a-x}\right)^{b-c}\left(\frac{b}{b-x}\right)^{c-a}\left(\frac{c}{c-x}\right)^{a-b}}{(a-b)(b-c)(c-a)} = -kt, \quad \text{where} \quad \begin{cases} x = 0 \\ t = 0. \end{cases}$$

8. Solve Ex. 7 if $a = b \neq c$; also when $a = b = c$. Note the very different forms of the solution in the three cases.

9. The rate at which water runs out of a tank through a small pipe issuing horizontally near the bottom of the tank is proportional to the square root of the height of the surface of the water above the pipe. If the tank is cylindrical and half empties in 30 min., show that it will completely empty in about 100 min.

10. Discuss Ex. 9 in case the tank were a right cone or frustum of a cone.

11. Consider a vertical column of air and assume that the pressure at any level is due to the weight of the air above. Show that $p = p_0 e^{-kh}$ gives the pressure at any height h, if Boyle's Law that the density of a gas varies as the pressure be used.

12. Work Ex. 11 under the assumption that the adiabatic law $p \propto \rho^{1.4}$ represents the conditions in the atmosphere. Show that in this case the pressure would become zero at a finite height. (If the proper numerical data are inserted, the height turns out to be about 20 miles. The adiabatic law seems to correspond better to the facts than Boyle's Law.)

13. Let l be the natural length of an elastic string, let Δl be the extension, and assume Hooke's Law that the force is proportional to the extension in the form $\Delta l = klF$. Let the string be held in a vertical position so as to elongate under its own weight W. Show that the elongation is $\frac{1}{2}kWl$.

14. The density of water under a pressure of p atmospheres is $\rho = 1 + 0.00004p$. Show that the surface of an ocean six miles deep is about 600 ft. below the position it would have if water were incompressible.

15. Show that the equations of the curve of equilibrium of a string or chain are

$$\frac{d}{ds}\left(T\frac{dr}{ds}\right) + \rho R = 0, \qquad \frac{d}{ds}\left(T\frac{rd\phi}{ds}\right) + \rho\Phi = 0$$

in polar coördinates, where R and Φ are the components of the force along the radius vector and perpendicular to it.

16. Show that $dT + \rho S ds = 0$ and $T + \rho RN = 0$ are the equations of equilibrium of a string if R is the radius of curvature and S and N are the tangential and normal components of the forces.

17.* Show that when a uniform chain is supported at two points and hangs down between the points under its own weight, the curve of equilibrium is the catenary.

18. Suppose the mass dm of the element ds of a chain is proportional to the projection dx of ds on the x-axis, and that the chain hangs in the field of gravity. Show that the curve is a parabola. (This is essentially the problem of the shape of the cables in a suspension bridge when the roadbed is of uniform linear density; for the weight of the cables is negligible compared to that of the roadbed.)

19. It is desired to string upon a cord a great many uniform heavy rods of varying lengths so that when the chord is hung up with the rods dangling from it the rods will be equally spaced along the horizontal and have their lower ends on the same level. Required the shape the chord will take. (It should be noted that the shape must be known before the rods can be cut in the proper lengths to hang as desired.) The weight of the chord may be neglected.

20. A masonry arch carries a horizontal roadbed. On the assumption that the material between the arch and the roadbed is of uniform density and that each element of the arch supports the weight of the material above it, find the shape of the arch.

21. In equations (4′) the integration may be carried through in terms of quadratures if ρY is a function of y alone; and similarly in Ex. 15 the integration may be carried through if $\Phi = 0$ and ρR is a function of r alone so that the field is central. Show that the results of thus carrying through the integration are the formulas

$$x + C' = \int \frac{C dy}{\sqrt{(\int \rho Y dy)^2 - C^2}}, \qquad \phi + C' = \int \frac{C dr/r}{\sqrt{(\int \rho R dr)^2 - C^2}}.$$

22. A particle falls from rest through the air, which is assumed to offer a resistance proportional to the velocity. Solve the problem with the initial conditions $v = 0$, $x = 0$, $t = 0$. Show that as the particle falls, the velocity does not increase indefinitely, but approaches a definite limit $V = mg/k$.

23. Solve Ex. 22 with the initial conditions $v = v_0$, $x = 0$, $t = 0$, where v_0 is greater than the limiting velocity V. Show that the particle slows down as it falls.

24. A particle rises through the air, which is assumed to resist proportionally to the square of the velocity. Solve the motion. Hyperbolic functions are useful.

25. Solve the problem analogous to Ex. 24 for a falling particle. Show that there is a limiting velocity $V = \sqrt{mg/k}$. If the particle were projected down with an initial velocity greater than V, it would slow down as in Ex. 23.

26. A particle falls towards a point which attracts it inversely as the square of the distance and directly as its mass. Find the relation between x and t and determine the total time T taken to reach the center. Initial conditions $v = 0$, $x = a$, $t = 0$.

$$\sqrt{\frac{2k}{a}}\, t = \frac{a}{2} \cos^{-1} \frac{2x - a}{a} + \sqrt{ax - x^2}, \qquad T = \pi k^{-\frac{1}{2}} \left(\frac{a}{2}\right)^{\frac{3}{2}}.$$

* Exercises 17–20 should be worked *ab initio* by the method by which (4) were derived, not by applying (4) directly.

27. A particle starts from the origin with a velocity V and moves in a medium which resists proportionally to the velocity. Find the relations between velocity and distance, velocity and time, and distance and time; also the limiting distance traversed.

$$v = V - kx/m, \qquad v = Ve^{-\frac{k}{m}t}, \qquad kx = mV(1 - e^{-\frac{k}{m}t}), \qquad mV/k.$$

28. Solve Ex. 27 under the assumption that the resistance varies as $\sqrt{v}$.

29. A particle falls toward a point which attracts inversely as the cube of the distance and directly as the mass. The initial conditions are $x = a$, $v = 0$, $t = 0$. Show that $x^2 = a^2 - kt^2/a^2$ and the total time of descent is $T = a^2/\sqrt{k}$.

30. A cylindrical spar buoy stands vertically in the water. The buoy is pressed down a little and released. Show that, if the resistance of the water and air be neglected, the motion is simple harmonic. Integrate and determine the constants from the initial conditions $x = 0$, $v = V$, $t = 0$, where x measures the displacement from the position of equilibrium.

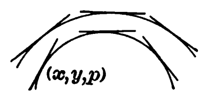
31. A particle slides down a rough inclined plane. Determine the motion. Note that of the force of gravity only the component $mg \sin i$ acts down the plane, whereas the component $mg \cos i$ acts perpendicularly to the plane and develops the force $\mu mg \cos i$ of friction. Here i is the inclination of the plane and μ is the coefficient of friction.

32. A bead is free to move upon a frictionless wire in the form of an inverted cycloid (vertex down). Show that the component of the weight along the tangent to the cycloid is proportional to the distance of the particle from the vertex. Hence determine the motion as simple harmonic and fix the constants of integration by the initial conditions that the particle starts from rest at the top of the cycloid.

33. Two equal weights are hanging at the end of an elastic string. One drops off. Determine completely the motion of the particle remaining.

34. One end of an elastic spring (such as is used in a spring balance) is attached rigidly to a point on a horizontal table. To the other end a particle is attached. If the particle be held at such a point that the spring is elongated by the amount a and then released, determine the motion on the assumption that the coefficient of friction between the particle and the table is μ; and discuss the possibility of different cases according as the force of friction is small or large relative to the force exerted by the spring.

85. Lineal element and differential equation. The idea of a curve as made up of the points upon it is familiar. Points, however, have no extension and therefore must be regarded not as pieces of a curve but merely as positions on it. Strictly speaking, the pieces of a curve are the elements Δs of arc; but for many purposes it is convenient to replace the complicated element Δs by a piece of the tangent to the curve at some point of the arc Δs, and from this point of view a curve is made up of an infinite number of infinitesimal elements tangent to it. This is analogous to the point of view by which a curve is regarded as made up of an infinite number of infinitesimal chords and is intimately related to the conception of the curve as the envelope of its tangents (§ 65). A point on a curve taken with an infinitesimal portion of the tangent to the curve at that point is called a *lineal element* of the curve. These concepts and definitions are clearly equally available in two or three dimensions. For the present the curves under discussion will be plane curves and the lineal elements will therefore all lie in a plane.



To specify any particular lineal element *three coördinates* x, y, p will be used, of which the two (x, y) determine the point through which the element passes and of which the third p is the slope of the element. If a curve $f(x, y) = 0$ is given, the slope at any point may be found by differentiation,

$$p = \frac{dy}{dx} = -\frac{\partial f}{\partial x} \Big/ \frac{\partial f}{\partial y}, \tag{6}$$

and hence the third coördinate p of the lineal elements of this particular curve is expressed in terms of the other two. If in place of one curve $f(x, y) = 0$ the whole family of curves $f(x, y) = C$, where C is an arbitrary constant, had been given, the slope p would still be found from (6), and it therefore appears that the third coördinate of the lineal elements of such a family of curves is expressible in terms of x and y.

In the more general case where the family of curves is given in the unsolved form $F(x, y, C) = 0$, the slope p is found by the same formula but it now depends apparently on C in addition to on x and y. If, however, the constant C be eliminated from the two equations

$$F(x, y, C) = 0 \quad \text{and} \quad \frac{\partial F}{\partial x} + \frac{\partial F}{\partial y} p = 0, \tag{7}$$

there will arise an equation $\Phi(x, y, p) = 0$ which connects the slope p of any curve of the family with the coördinates (x, y) of any point through which a curve of the family passes and at which the slope of that curve is p. Hence it appears that the three coördinates (x, y, p) of the lineal elements of all the curves of a family are connected by an equation $\Phi(x, y, p) = 0$, just as the coördinates (x, y, z) of the points of a surface are connected by an equation $\Phi(x, y, z) = 0$. As the equation $\Phi(x, y, z) = 0$ is called the equation of the surface, so the equation $\Phi(x, y, p) = 0$ is called the equation of the family of curves; it is, however, not the finite equation $F(x, y, C) = 0$ but the differential equation of the family, because it involves the derivative $p = dy/dx$ of y by x instead of the parameter C.

As an example of the elimination of a constant, consider the case of the parabolas

$$y^2 = Cx \quad \text{or} \quad y^2/x = C.$$

The differentiation of the equation in the second form gives at once

$$-y^2/x^2 + 2yp/x = 0 \quad \text{or} \quad y = 2xp$$

as the differential equation of the family. In the unsolved form the work is

$$2yp = C, \qquad y^2 = 2ypx, \qquad y = 2xp.$$

The result is, of course, the same in either case. For the family here treated it makes little difference which method is followed. As a general rule it is perhaps best to solve for the constant if the solution is simple and leads to a simple form of the function $f(x, y)$; whereas if the solution is not simple or the form of the function is complicated, it is best to differentiate first because the differentiated equation may be simpler to solve for the constant than the original equation, or because the elimination of the constant between the two equations can be conducted advantageously.

If an equation $\Phi(x, y, p) = 0$ connecting the three coördinates of the lineal element be given, the elements which satisfy the equation may be plotted much as a surface is plotted; that is, a pair of values (x, y) may be assumed and substituted in the equation, the equation may then be solved for one or more values of p, and lineal elements with these values of p may be drawn through the point (x, y). In this manner the elements through as many points as desired may be found. The detached elements are of interest and significance chiefly from the fact that they can be *assembled into curves*, — in fact, into the curves of a family $F(x, y, C) = 0$ of which the equation $\Phi(x, y, p) = 0$ is the differential equation. This is the converse of the problem treated above and requires the integration of the differential equation $\Phi(x, y, p) = 0$ for its solution. In some simple cases the assembling may be accomplished intuitively from the geometric properties implied in the equation, in other cases it follows from the integration of the equation by analytic means, in other cases it can be done only approximately and by methods of computation.

As an example of intuitively assembling the lineal elements into curves, take

$$\Phi(x, y, p) = y^2p^2 + y^2 - r^2 = 0 \quad \text{or} \quad p = \pm\frac{\sqrt{r^2 - y^2}}{y}.$$

The quantity $\sqrt{r^2 - y^2}$ may be interpreted as one leg of a right triangle of which y is the other leg and r the hypotenuse. The slope of the hypotenuse is then $\pm y/\sqrt{r^2 - y^2}$ according to the position of the figure, and the differential equation $\Phi(x, y, p) = 0$ states that the coördinate p of the lineal element which satisfies it is the negative reciprocal of this slope. Hence the lineal element is perpendicular to the hypotenuse. It therefore appears that the lineal elements are tangent to circles of radius r described about points of the x-axis. The equation of these circles is

$(x-C)^2+y^2=r^2$, and this is therefore the integral of the differential equation. The correctness of this integral may be checked by direct integration. For

$$p=\frac{dy}{dx}=\pm\frac{\sqrt{r^2-y^2}}{y} \quad \text{or} \quad \frac{ydy}{\sqrt{r^2-y^2}}=dx \quad \text{or} \quad \sqrt{r^2-y^2}=x-C.$$

86. In geometric problems which relate the slope of the tangent of a curve to other lines in the figure, it is clear that not the tangent but the lineal element is the vital thing. Among such problems that of the *orthogonal trajectories* (or trajectories under any angle) of a given family of curves is of especial importance. If two families of curves are so related that the angle at which any curve of one of the families cuts any curve of the other family is a right angle, then the curves of either family are said to be the orthogonal trajectories of the curves of the other family. Hence at any point (x, y) at which two curves belonging to the different families intersect, there are two lineal elements, one belonging to each curve, which are perpendicular. As the slopes of two perpendicular lines are the negative reciprocals of each other, it follows that if the coördinates of one lineal element are (x, y, p) the coördinates of the other are $(x, y, -1/p)$; and if the coördinates of the lineal element (x, y, p) satisfy the equation $\Phi(x, y, p)=0$, the coördinates of the orthogonal lineal element must satisfy $\Phi(x, y, -1/p)=0$. Therefore the *rule for finding the orthogonal trajectories of the curves $F(x, y, C)=0$ is to find first the differential equation $\Phi(x, y, p)=0$ of the family, then to replace p by $-1/p$ to find the differential equation of the orthogonal family, and finally to integrate this equation to find the family.* It may be noted that if $F(z)=X(x, y)+iY(x, y)$ is a function of $z=x+iy$ (§ 73), the families $X(x, y)=C$ and $Y(x, y)=K$ are orthogonal.

As a problem in orthogonal trajectories find the trajectories of the semicubical parabolas $(x-C)^3=y^2$. The differential equation of this family is found as

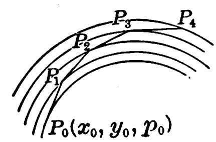
$$3(x-C)^2=2yp, \qquad x-C=(\tfrac{2}{3}yp)^{\frac{1}{2}}, \qquad (\tfrac{2}{3}yp)^{\frac{3}{2}}=y^2 \quad \text{or} \quad \tfrac{2}{3}p=y^{\frac{1}{3}}.$$

This is the differential equation of the given family. Replace p by $-1/p$ and integrate:

$$-\frac{2}{3p}=y^{\frac{1}{3}} \quad \text{or} \quad 1+\frac{3}{2}py^{\frac{1}{3}}=0 \quad \text{or} \quad dx+\frac{3}{2}y^{\frac{1}{3}}dy=0, \quad \text{and} \quad x+\frac{9}{8}y^{\frac{4}{3}}=C.$$

Thus the differential equation and finite equation of the orthogonal family are found. The curves look something like parabolas with axis horizontal and vertex toward the right.

Given a differential equation $\Phi(x, y, p)=0$ or, in solved form, $p=\phi(x, y)$; the *lineal element affords a means for obtaining graphically and numerically an approximation to the solution which passes through* *an assigned point* $P_0(x_0, y_0)$. For the value p_0 of p at this point may be computed from the equation and a lineal element P_0P_1 may be drawn, the length being taken small. As the lineal element is tangent to the curve, its end point will not lie upon the curve but will depart from it by an infinitesimal of higher order. Next the slope p_1 of the lineal element which satisfies the equation and passes through P_1 may be found and the element P_1P_2 may be drawn. This element will not be tangent to the desired solution but to a solution lying near that one. Next the element P_2P_3 may be drawn, and so on. The broken line $P_0P_1P_2P_3\cdots$ is clearly an approximation to the solution and will be a better approximation the shorter the elements P_iP_{i+1} are taken. If the radius of curvature of the solution at P_0 is not great, the curve will be bending rapidly and the elements must be taken fairly short in order to get a fair approximation; but if the radius of curvature is great, the elements need not be taken so small. (This method of approximate graphical solution indicates a method which is of value in proving by the method of limits that the equation $p = \phi(x, y)$ actually has a solution; but that matter will not be treated here.)



Let it be required to plot approximately that solution of $yp + x = 0$ which passes through (0, 1) and thus to find the ordinate for $x = 0.5$, and the area under the curve and the length of the curve to this point. Instead of assuming the lengths of the successive lineal elements, let the lengths of successive increments δx of x be taken as $\delta x = 0.1$. At the start $x_0 = 0$, $y_0 = 1$, and from $p = -x/y$ it follows that $p_0 = 0$. The increment δy of y acquired in moving along the tangent is $\delta y = p\delta x = 0$. Hence the new point of departure (x_1, y_1) is (0.1, 1) and the new slope is $p_1 = -x_1/y_1 = -0.1$. The results of the work, as it is continued, may be grouped in the table. Hence it appears that the final ordinate is $y = 0.90$. By adding up the trapezoids the area is computed as 0.48, and by finding the elements $\delta s = \sqrt{\delta x^2 + \delta y^2}$ the length is found as 0.51. Now the particular equation here treated can be integrated.

i	δx	δy	x_i	y_i	p_i
0	...	...	0.	1.00	0.
1	0.1	0.	0.1	1.00	− 0.1
2	0.1	− 0.01	0.2	0.99	− 0.2
3	0.1	− 0.02	0.3	0.97	− 0.31
4	0.1	− 0.03	0.4	0.94	− 0.43
5	0.1	− 0.04	0.5	0.90	...

$$yp + x = 0, \qquad ydy + xdx = 0, \qquad x^2 + y^2 = C, \quad \text{and hence} \quad x^2 + y^2 = 1$$

is the solution which passes through (0, 1). The ordinate, area, and length found from the curve are therefore 0.87, 0.48, 0.52 respectively. The errors in the approximate results to two places are therefore respectively 3, 0, 2 per cent. If δx had been chosen as 0.01 and four places had been kept in the computations, the errors would have been smaller.

EXERCISES

1. In the following cases eliminate the constant C to find the differential equation of the family given:

(α) $x^2 = 2Cy + C^2$, (β) $y = Cx + \sqrt{1 - C^2}$,

(γ) $x^2 - y^2 = Cx$, (δ) $y = x\tan(x + C)$,

(ϵ) $\dfrac{x^2}{a^2 - C} + \dfrac{y^2}{b^2 - C} = 1$, *Ans.* $\left(\dfrac{dy}{dx}\right)^2 + \dfrac{(x^2 - y^2) - (a^2 - b^2)}{xy}\dfrac{dy}{dx} - 1 = 0.$

2. Plot the lineal elements and intuitively assemble them into the solution:

$$(\alpha)\ yp + x = 0, \qquad (\beta)\ xp - y = 0, \qquad (\gamma)\ r\frac{d\phi}{dr} = 1.$$

Check the results by direct integration of the differential equations.

3. Lines drawn from the points $(\pm c, 0)$ to the lineal element are equally inclined to it. Show that the differential equation is that of Ex. 1 (ϵ). What are the curves?

4. The trapezoidal area under the lineal element equals the sectorial area formed by joining the origin to the extremities of the element (disregarding infinitesimals of higher order). (α) Find the differential equation and integrate. (β) Solve the same problem where the areas are equal in magnitude but opposite in sign. What are the curves?

5. Find the orthogonal trajectories of the following families. Sketch the curves.

(α) parabolas $y^2 = 2Cx$, *Ans.* ellipses $2x^2 + y^2 = C$.

(β) exponentials $y = Ce^{kx}$, *Ans.* parabolas $\frac{1}{2}ky^2 + x = C$.

(γ) circles $(x - C)^2 + y^2 = a^2$, *Ans.* tractrices.

(δ) $x^2 - y^2 = C^2$, (ϵ) $Cy^2 = x^3$, (ζ) $x^{\frac{2}{3}} + y^{\frac{2}{3}} = C^{\frac{2}{3}}$.

6. Show from the answer to Ex. 1 (ϵ) that the family is self-orthogonal and illustrate with a sketch. From the fact that the lineal element of a parabola makes equal angles with the axis and with the line drawn to the focus, derive the differential equation of all coaxial confocal parabolas and show that the family is self-orthogonal.

7. If $\Phi(x, y, p) = 0$ is the differential equation of a family, show

$$\Phi\left(x, y, \frac{p - m}{1 + mp}\right) = 0 \quad \text{and} \quad \Phi\left(x, y, \frac{p + m}{1 - mp}\right) = 0$$

are the differential equations of the family whose curves cut those of the given family at $\tan^{-1} m$. What is the difference between these two cases?

8. Show that the differential equations

$$\Phi\left(\frac{dr}{d\phi}, r, \phi\right) = 0 \quad \text{and} \quad \Phi\left(-r^2\frac{d\phi}{dr}, r, \phi\right) = 0$$

define orthogonal families in polar coördinates, and write the equation of the family which cuts the first of these at the constant angle $\tan^{-1} m$.

9. Find the orthogonal trajectories of the following families. Sketch.

(α) $r = e^{C\phi}$, (β) $r = C(1 - \cos\phi)$, (γ) $r = C\phi$, (δ) $r^2 = C^2\cos 2\phi$.

10. Recompute the approximate solution of $yp + x = 0$ under the conditions of the text but with $\delta x = 0.05$, and carry the work to three decimals.

11. Plot the approximate solution of $p = xy$ between (1, 1) and the y-axis. Take $\delta x = -0.2$. Find the ordinate, area, and length. Check by integration and comparison.

12. Plot the approximate solution of $p = -x$ through (1, 1), taking $\delta x = 0.1$ and following the curve to its intersection with the x-axis. Find also the area and the length.

13. Plot the solution of $p = \sqrt{x^2 + y^2}$ from the point (0, 1) to its intersection with the x-axis. Take $\delta x = -0.2$ and find the area and length.

14. Plot the solution of $p = s$ which starts from the origin into the first quadrant (s is the length of the arc). Take $\delta x = 0.1$ and carry the work for five steps to find the final ordinate, the area, and the length. Compare with the true integral.

87. The higher derivatives; analytic approximations. Although a differential equation $\Phi(x, y, y') = 0$ does not determine the relation between x and y without the application of some process equivalent to integration, it does afford a means of computing the higher derivatives simply by differentiation. Thus

$$\frac{d\Phi}{dx} = \frac{\partial\Phi}{\partial x} + \frac{\partial\Phi}{\partial y}y' + \frac{\partial\Phi}{\partial y'}y'' = 0$$

is an equation which may be solved for y'' as a function of x, y, y'; and y'' may therefore be expressed in terms of x and y by means of $\Phi(x, y, y') = 0$. A further differentiation gives the equation

$$\frac{d^2\Phi}{dx^2} = \frac{\partial^2\Phi}{\partial x^2} + 2\frac{\partial^2\Phi}{\partial x\partial y}y' + 2\frac{\partial^2\Phi}{\partial x\partial y'}y'' + \frac{\partial^2\Phi}{\partial y^2}y'^2 + 2\frac{\partial^2\Phi}{\partial y\partial y'}y'y''$$
$$+ \frac{\partial^2\Phi}{\partial y'^2}y''^2 + \frac{\partial\Phi}{\partial y}y'' + \frac{\partial\Phi}{\partial y'}y''' = 0,$$

which may be solved for y''' in terms of x, y, y', y''; and hence, by the preceding results, y''' is expressible as a function of x and y; and so on to all the higher derivatives. In this way any property of the integrals of $\Phi(x, y, y') = 0$ which, like the radius of curvature, is expressible in terms of the derivatives, may be found as a function of x and y.

As the differential equation $\Phi(x, y, y') = 0$ defines y' and all the higher derivatives as functions of x, y, it is clear that the values of the derivatives may be found as y_0', y_0'', y_0''', $\cdots$ at any given point (x_0, y_0). Hence it is possible to write the series

$$y = y_0 + y_0'(x - x_0) + \tfrac{1}{2}y_0''(x - x_0)^2 + \tfrac{1}{6}y_0'''(x - x_0)^3 + \cdots. \qquad (8)$$

If this power series in $x - x_0$ converges, it defines y as a function of x for values of x near x_0; it is indeed *the Taylor development of the* *function* y (§ 167). The convergence is assumed. Then

$$y' = y_0' + y_0''(x - x_0) + \tfrac{1}{2} y_0'''(x - x_0)^2 + \cdots.$$

It may be shown that the function y defined by the series actually satisfies the differential equation $\Phi(x, y, y') = 0$, that is, that

$$\Omega(x) = \Phi[x, y_0 + y_0'(x - x_0) + \tfrac{1}{2} y_0''(x - x_0)^2 + \cdots, y_0' + y_0''(x - x_0) + \cdots] = 0$$

for all values of x near x_0. To prove this accurately, however, is beyond the scope of the present discussion; the fact may be taken for granted. Hence an analytic expansion for the integral of a differential equation has been found.

As an example of computation with higher derivatives let it be required to determine the radius of curvature of that solution of $y' = \tan(y/x)$ which passes through (1, 1). Here the slope $y'_{(1,1)}$ at (1, 1) is $\tan 1 = 1.557$. The second derivative is

$$y'' = \frac{dy'}{dx} = \frac{d}{dx}\tan\frac{y}{x} = \sec^2\frac{y}{x}\,\frac{xy' - y}{x^2}.$$

From these data the radius of curvature is found to be

$$R = \frac{(1 + y'^2)^{\frac{3}{2}}}{y''} = \sec\frac{y}{x}\,\frac{x^2}{xy' - y}, \qquad R_{(1,1)} = \sec 1\,\frac{1}{\tan 1 - 1} = 3.250.$$

The equation of the circle of curvature may also be found. For as $y''_{(1,1)}$ is positive, the curve is concave up. Hence $(1 - 3.250 \sin 1,\ 1 + 3.250 \cos 1)$ is the center of curvature; and the circle is

$$(x + 1.735)^2 + (y - 2.757)^2 = (3.250)^2.$$

As a second example let four terms of the expansion of that integral of $x \tan y' = y$ which passes through (2, 1) be found. The differential equation may be solved; then

$$\frac{dy}{dx} = \tan^{-1}\left(\frac{y}{x}\right), \qquad \frac{d^2y}{dx^2} = \frac{xy' - y}{x^2 + y^2},$$

$$\frac{d^3y}{dx^3} = \frac{(x^2 + y^2)(x - 1)y'' + (3y^2 - x^2)y' - 2xyy'^2 + 2xy}{(x^2 + y^2)^2}.$$

Now it must be noted that the problem is not wholly determinate; for y' is multiple valued and any one of the values for $\tan^{-1}\frac{1}{2}$ may be taken as the slope of a solution through (2, 1). Suppose that the angle be taken in the first quadrant; then $\tan^{-1}\frac{1}{2} = 0.462$. Substituting this in y'', we find $y''_{(2,1)} = -0.0152$; and hence may be found $y'''_{(2,1)} = 0.110$. The series for y to four terms is therefore

$$y = 1 + 0.462(x - 2) - 0.0076(x - 2)^2 + 0.018(x - 3)^3.$$

It may be noted that it is generally simpler not to express the higher derivatives in terms of x and y, but to compute each one successively from the preceding ones.

88. Picard has given a method for the integration of the equation $y' = \phi(x, y)$ by *successive approximations* which, although of the highest theoretic value and importance, is not particularly suitable to analytic uses in finding an approximate solution. The method is this. Let the equation $y' = \phi(x, y)$ be given in solved form, and suppose (x_0, y_0) is the point through which the solution is to pass. To find the first approximation let y be held constant and equal to y_0, and integrate the equation $y' = \phi(x, y_0)$. Thus

$$dy = \phi(x, y_0)\,dx; \qquad y = y_0 + \int_{x_0}^{x} \phi(x, y_0)\,dx = f_1(x), \tag{9}$$

where it will be noticed that the constant of integration has been chosen so that the curve passes through (x_0, y_0). For the second approximation let y have the value just found, substitute this in $\phi(x, y)$, and integrate again. Then

$$y = y_0 + \int_{x_0}^{x} \phi\left[x, y_0 + \int_{x_0}^{x} \phi(x, y_0)\,dx\right] dx = f_2(x). \tag{9'}$$

With this new value for y continue as before. The successive determinations of y as a function of x actually converge toward a limiting function which is a solution of the equation and which passes through (x_0, y_0). It may be noted that at each step of the work an integration is required. The difficulty of actually performing this integration in formal practice limits the usefulness of the method in such cases. It is clear, however, that with an integrating machine such as the integraph the method could be applied as rapidly as the curves $\phi(x, f_i(x))$ could be plotted.

To see how the method works, consider the integration of $y' = x + y$ to find the integral through (1, 1). For the first approximation $y = 1$. Then

$$dy = (x + 1)\,dx, \qquad y = \tfrac{1}{2}x^2 + x + C, \qquad y = \tfrac{1}{2}x^2 + x - \tfrac{1}{2} = f_1(x).$$

From this value of y the next approximation may be found, and then still another:

$$dy = [x + (\tfrac{1}{2}x^2 + x - \tfrac{1}{2})]\,dx, \qquad y = \tfrac{1}{6}x^3 + x^2 - \tfrac{1}{2}x + \tfrac{1}{3} = f_2(x),$$
$$dy = [x + f_2(x)]\,dx, \qquad y = \tfrac{1}{24}x^4 + \tfrac{1}{3}x^3 + \tfrac{1}{4}x^2 + \tfrac{1}{3}x + \tfrac{1}{24}.$$

In this case there are no difficulties which would prevent any number of applications of the method. In fact it is evident that if y' is a polynomial in x and y, the result of any number of applications of the method will be a polynomial in x.

The method of *undetermined coefficients* may often be employed to advantage to develop the solution of a differential equation into a series. The result is of course identical with that obtained by the application of successive differentiation and Taylor's series as above; the work is sometimes shorter. Let the equation be in the form $y' = \phi(x, y)$ and assume an integral in the form

$$y = y_0 + a_1(x - x_0) + a_2(x - x_0)^2 + a_3(x - x_0)^3 + \cdots. \tag{10}$$

Then $\phi(x, y)$ may also be expanded into a series, say,

$$\phi(x, y) = A_0 + A_1(x - x_0) + A_2(x - x_0)^2 + A_3(x - x_0)^3 + \cdots.$$

But by differentiating the assumed form for y we have

$$y' = a_1 + 2a_2(x - x_0) + 3a_3(x - x_0)^2 + 4a_4(x - x_0)^3 + \cdots.$$

Thus there arise two different expressions as series in $x - x_0$ for the function y', and therefore the corresponding coefficients must be equal. The resulting set of equations

$$a_1 = A_0, \quad 2a_2 = A_1, \quad 3a_3 = A_2, \quad 4a_4 = A_3, \quad \cdots \tag{11}$$

may be solved successively for the undetermined coefficients a_1, a_2, a_3, $a_4, \cdots$ which enter into the assumed expansion. This method is particularly useful when the form of the differential equation is such that some of the terms may be omitted from the assumed expansion (see Ex. 14).

As an example in the use of undetermined coefficients consider that solution of the equation $y' = \sqrt{x^2 + 3y^2}$ which passes through (1, 1). The expansion will proceed according to powers of $x - 1$, and for convenience the variable may be changed to $t = x - 1$ so that

$$\frac{dy}{dt} = \sqrt{(t+1)^2 + 3y^2}, \qquad y = 1 + a_1 t + a_2 t^2 + a_3 t^3 + a_4 t^4 + \cdots$$

are the equation and the assumed expansion. One expression for y' is

$$y' = a_1 + 2a_2 t + 3a_3 t^2 + 4a_4 t^3 + \cdots.$$

To find the other it is necessary to expand into a series in t the expression

$$y' = \sqrt{(1+t)^2 + 3(1 + a_1 t + a_2 t^2 + a_3 t^3)^2}.$$

If this had to be done by Maclaurin's series, nothing would be gained over the method of § 87; but in this and many other cases algebraic methods and known expansions may be applied (§ 32). First square y and retain only terms up to the third power. Hence

$$y' = 2\sqrt{1 + \tfrac{1}{2}(1 + 3a_1)t + \tfrac{1}{4}(1 + 6a_2 + 3a_1^2)t^2 + \tfrac{3}{2}(a_1 a_2 + a_3)t^3}.$$

Now let the quantity under the radical be called $1 + h$ and expand so that

$$y' = 2\sqrt{1+h} = 2(1 + \tfrac{1}{2}h - \tfrac{1}{8}h^2 + \tfrac{1}{16}h^3).$$

Finally raise h to the indicated powers and collect in powers of t. Then

$$y' = 2 + \left.\begin{array}{r} t \\ \tfrac{1}{2}(1 + 3a_1) \\ \\ \end{array}\right| \left.\begin{array}{l} \hfill t^2 \\ + \tfrac{1}{4}(1 + 6a_2 + 3a_1^2) \\ - \tfrac{1}{16}(1 + 3a_1)^2 \\ \end{array}\right| \left.\begin{array}{l} \hfill t^3 \\ + \tfrac{3}{2}(a_1 a_2 + a_3) \\ - \tfrac{1}{16}(1 + 3a_1)(1 + 6a_2 + 3a_1^2) \\ + \tfrac{1}{64}(1 + 3a_1)^3 \end{array}\right|.$$

Hence the successive equations for determining the coefficients are $a_1 = 2$ and

$$2a_2 = \tfrac{1}{2}(1 + 3a_1) \text{ or } a_2 = \tfrac{7}{4},$$
$$3a_3 = \tfrac{1}{4}(1 + 6a_2 + 3a_1^2) - \tfrac{1}{16}(1 + 3a_1)^2 \text{ or } a_3 = \tfrac{15}{16},$$
$$4a_4 = \tfrac{3}{2}(a_1a_2 + a_3) - \tfrac{1}{16}(1 + 3a_1)(1 + 6a_2 + 3a_1^2) + \tfrac{1}{64}(1 + 3a_1)^3 \text{ or } a_4 = \tfrac{111}{256}.$$

Therefore to five terms the expansion desired is

$$y = 1 + 2(x - 1) + \tfrac{7}{4}(x - 1)^2 + \tfrac{15}{16}(x - 1)^3 + \tfrac{111}{256}(x - 1)^4.$$

The methods of developing a solution by Taylor's series or by undetermined coefficients apply equally well to equations of higher order. For example consider an equation of the second order in solved form $y'' = \phi(x, y, y')$ and its derivatives

$$y''' = \frac{\partial\phi}{\partial x} + \frac{\partial\phi}{\partial y}y' + \frac{\partial\phi}{\partial y'}y''$$

$$y^{\text{iv}} = \frac{\partial^2\phi}{\partial x^2} + 2\frac{\partial^2\phi}{\partial x\partial y}y' + 2\frac{\partial^2\phi}{\partial x\partial y'}y'' + \frac{\partial^2\phi}{\partial y^2}y'^2 + 2\frac{\partial^2\phi}{\partial y\partial y'}y'y'' + \frac{\partial^2\phi}{\partial y'^2}y''^2 + \frac{\partial\phi}{\partial y}y'' + \frac{\partial\phi}{\partial y'}y'''.$$

Evidently the higher derivatives of y may be obtained in terms of x, y, y'; and y itself may be written in the expanded form

$$y = y_0 + y_0'(x - x_0) + \tfrac{1}{2}y_0''(x - x_0)^2 + \tfrac{1}{6}y_0'''(x - x_0)^3 + \tfrac{1}{24}y_0^{\text{iv}}(x - x_0)^4 + \cdots, \tag{12}$$

where any desired values may be attributed to the ordinate y_0 at which the curve cuts the line $x = x_0$, and to the slope y_0' of the curve at that point. Moreover the coefficients y_0'', y_0''', $\cdots$ are determined in such a way that they depend on the assumed values of y_0 and y_0'. It therefore is seen that the solution (12) of the differential equation of the second order really involves two arbitrary constants, and the justification of writing it as $F(x, y, C_1, C_2) = 0$ is clear.

In following out the method of undetermined coefficients a solution of the equation would be assumed in the form

$$y = y_0 + y_0'(x - x_0) + a_2(x - x_0)^2 + a_3(x - x_0)^3 + a_4(x - x_0)^4 + \cdots, \tag{13}$$

from which y' and y'' would be obtained by differentiation. Then if the series for y and y' be substituted in $y'' = \phi(x, y, y')$ and the result arranged as a series, a second expression for y'' is obtained and the comparison of the coefficients in the two series will afford a set of equations from which the successive coefficients may be found in terms of y_0 and y_0' by solution. These results may clearly be generalized to the case of differential equations of the nth order, whereof the solutions will depend on n arbitrary constants, namely, the values assumed for y and its first $n - 1$ derivatives when $x = x_0$.

EXERCISES

1. Find the radii and circles of curvature of the solutions of the following equations at the points indicated:

(α) $y' = \sqrt{x^2 + y^2}$ at $(0, 1)$, (β) $yy' + x = 0$ at (x_0, y_0).

2. Find $y'''_{(1, 1)} = (5\sqrt{2} - 2)/4$ if $y' = \sqrt{x^2 + y^2}$.

3. Given the equation $y^2y'^3 + xyy'^2 - yy' + x^2 = 0$ of the third degree in y' so that there will be three solutions with different slopes through any ordinary point (x, y). Find the radii of curvature of the three solutions through $(0, 1)$.

4. Find three terms in the expansion of the solution of $y' = e^{xy}$ about $(2, \frac{1}{2})$.

5. Find four terms in the expansion of the solution of $y = \log \sin xy$ about $(\frac{1}{2}\pi, 1)$.

6. Expand the solution of $y' = xy$ about $(1, y_0)$ to five terms.

7. Expand the solution of $y' = \tan(y/x)$ about $(1, 0)$ to four terms. Note that here x should be expanded in terms of y, not y in terms of x.

8. Expand two of the solutions of $y^2y'^3 + xyy'^2 - yy' + x^2 = 0$ about $(-2, 1)$ to four terms.

9. Obtain four successive approximations to the integral of $y' = xy$ through $(1, 1)$.

10. Find four successive approximations to the integral of $y' = x + y$ through $(0, y_0)$.

11. Show by successive approximations that the integral of $y' = y$ through $(0, y_0)$ is the well-known $y = y_0e^x$.

12. Carry the approximations to the solution of $y' = -x/y$ through $(0, 1)$ as far as you can integrate, and plot each approximation on the same figure with the exact integral.

13. Find by the method of undetermined coefficients the number of terms indicated in the expansions of the solutions of these differential equations about the points given:

(α) $y' = \sqrt{x + y}$, five terms, $(0, 1)$, (β) $y' = \sqrt{x + y}$, four terms, $(1, 3)$,

(γ) $y' = x + y$, n terms, $(0, y_0)$, (δ) $y' = \sqrt{x^2 + y^2}$, four terms, $(\frac{3}{8}, \frac{1}{4})$.

14. If the solution of an equation is to be expanded about $(0, y_0)$ and if the change of x into $-x$ and y' into $-y'$ does not alter the equation, the solution is necessarily symmetric with respect to the y-axis and the expansion may be assumed to contain only even powers of x. If the solution is to be expanded about $(0, 0)$ and a change of x into $-x$ and y into $-y$ does not alter the equation, the solution is symmetric with respect to the origin and the expansion may be assumed in odd powers. Obtain the expansions to four terms in the following cases and compare the labor involved in the method of undetermined coefficients with that which would be involved in performing the requisite six or seven differentiations for the application of Maclaurin's series:

(α) $y' = \dfrac{x}{\sqrt{x^2 + y^2}}$ about $(0, 2)$, (β) $y' = \sin xy$ about $(0, 1)$,

(γ) $y' = e^{xy}$ about $(0, 0)$, (δ) $y' = x^3y + xy^3$ about $(0, 0)$.

15. Expand to and including the term x^4:

(α) $y'' = y'^2 + xy$ about $x_0 = 0$, $y_0 = a_0$, $y'_0 = a_1$ (by both methods),

(β) $xy'' + y' + y = 0$ about $x_0 = 0$, $y_0 = a_0$, $y'_0 = -a_0$ (by und. coeffs.).

CHAPTER VIII

THE COMMONER ORDINARY DIFFERENTIAL EQUATIONS

89. Integration by separating the variables. If a differential equation of the first order may be solved for y' so that

$$y' = \phi(x, y) \quad \text{or} \quad M(x, y)\,dx + N(x, y)\,dy = 0 \tag{1}$$

(where the functions ϕ, M, N are single valued or where only one specific branch of each function is selected in case the solution leads to multiple valued functions), the differential equation involves only the first power of the derivative and is said to be of the first degree. If, furthermore, it so happens that the functions ϕ, M, N are products of functions of x and functions of y so that the equation (1) takes the form

$$y' = \phi_1(x)\,\phi_2(y) \quad \text{or} \quad M_1(x)M_2(y)\,dx + N_1(x)\,N_2(y)\,dy = 0, \tag{2}$$

it is clear that the variables may be separated in the manner

$$\frac{dy}{\phi_2(y)} = \phi_1(x)\,dx \quad \text{or} \quad \frac{M_1(x)}{N_1(x)}\,dx + \frac{N_2(y)}{M_2(y)}\,dy = 0, \tag{2'}$$

and the integration is then immediately performed by integrating each side of the equation. It was in this way that the numerous problems considered in Chap. VII were solved.

As an example consider the equation $yy' + xy^2 = x$. Here

$$ydy + x(y^2 - 1)\,dx = 0 \quad \text{or} \quad \frac{ydy}{y^2 - 1} + xdx = 0,$$

and $$\tfrac{1}{2}\log(y^2 - 1) + \tfrac{1}{2}x^2 = C \quad \text{or} \quad (y^2 - 1)\,e^{x^2} = C.$$

The second form of the solution is found by taking the exponential of both sides of the first form after multiplying by 2.

In some differential equations (1) in which the variables are not immediately separable as above, the introduction of some change of variable, whether of the dependent or independent variable or both, may lead to a differential equation in which the new variables are separated and the integration may be accomplished. The selection of the proper change of variable is in general a matter for the exercise of ingenuity; succeeding paragraphs, however, will point out some special types of equations for which a definite type of substitution is known to accomplish the separation.

As an example consider the equation $xdy - ydx = x\sqrt{x^2 + y^2}\,dx$, where the variables are clearly not separable without substitution. The presence of $\sqrt{x^2 + y^2}$ suggests a change to polar coördinates. The work of finding the solution is:

$$x = r\cos\theta,\quad y = r\sin\theta,\quad dx = \cos\theta dr - r\sin\theta d\theta,\quad dy = \sin\theta dr + r\cos\theta d\theta;$$

then $$xdy - ydx = r^2 d\theta,\qquad x\sqrt{x^2 + y^2}\,dx = r^2\cos\theta d(r\cos\theta).$$

Hence the differential equation may be written in the form

$$r^2 d\theta = r^2\cos\theta d(r\cos\theta)\quad\text{or}\quad \sec\theta d\theta = d(r\cos\theta),$$

and $$\log\tan(\tfrac{1}{2}\theta + \tfrac{1}{4}\pi) = r\cos\theta + C\quad\text{or}\quad \log\frac{1+\sin\theta}{\cos\theta} = x + C.$$

Hence $$\frac{\sqrt{x^2 + y^2} + y}{x} = Ce^x\qquad\text{(on substitution for }\theta).$$

Another change of variable which works, is to let $y = vx$. Then the work is:

$$x(vdx + xdv) - vxdx = x^2\sqrt{1 + v^2}\,dx\ \text{or}\ dv = \sqrt{1 + v^2}\,dx.$$

Then $$\frac{dv}{\sqrt{1 + v^2}} = dx,\qquad \sinh^{-1}v = x + C,\qquad y = x\sinh(x + C).$$

This solution turns out to be shorter and the answer appears in neater form than before obtained. The great difference of form that may arise in the answer when different methods of integration are employed, is a noteworthy fact, and renders a set of answers practically worthless; two solvers may frequently waste more time in trying to get their answers reduced to a common form than each would spend in solving the problem in two ways.

90. If in the equation $y' = \phi(x, y)$ the function ϕ turns out to be $\phi(y/x)$, a function of y/x alone, that is, if the functions M and N are homogeneous functions of x, y and of the same order (§ 53), the differential equation is said to be *homogeneous* and the change of variable $y = vx$ or $x = vy$ will always result in separating the variables. The statement may be tabulated as:

$$\text{if}\qquad \frac{dy}{dx} = \phi\left(\frac{y}{x}\right),\qquad \text{substitute}\begin{cases} y = vx \\ \text{or } x = vy. \end{cases}\tag{3}$$

A sort of corollary case is given in Ex. 6 below.

As an example take $y\left(1 + e^{\frac{x}{y}}\right)dx + e^{\frac{x}{y}}(y - x)\,dy = 0$, of which the homogeneity is perhaps somewhat disguised. Here it is better to choose $x = vy$. Then

$$(1 + e^v)\,dx + e^v(1 - v)\,dy = 0\quad\text{and}\quad dx = vdy + ydv.$$

Hence $$(v + e^v)\,dy + y(1 + e^v)\,dv = 0\quad\text{or}\quad \frac{dy}{y} + \frac{1 + e^v}{v + e^v}dy = 0.$$

Hence $$\log y + \log(v + e^v) = C\quad\text{or}\quad x + ye^{\frac{x}{y}} = C.$$

If the differential equation may be arranged so that

$$\frac{dy}{dx} + X_1(x)\,y = X_2(x)\,y^n \quad \text{or} \quad \frac{dx}{dy} + Y_1(y)\,x = Y_2(y)\,x^n, \tag{4}$$

where the second form differs from the first only through the interchange of x and y and where X_1 and X_2 are functions of x alone and Y_1 and Y_2 functions of y, the equation is called a *Bernoulli equation;* and in particular if $n = 0$, so that the dependent variable does not occur on the right-hand side, the equation is called *linear.* The substitution which separates the variables in the respective cases is

$$y = ve^{-\int X_1(x)\,dx} \quad \text{or} \quad x = ve^{-\int Y_1(y)\,dy}. \tag{5}$$

To show that the separation is really accomplished and to find a general formula for the solution of any Bernoulli or linear equation, the substitution may be carried out formally. For

$$\frac{dy}{dx} = \frac{dv}{dx}\,e^{-\int X_1 dx} - vX_1 e^{-\int X dx}.$$

The substitution of this value in the equation gives

$$\frac{dv}{dx}\,e^{-\int X_1 dx} = X_2 v^n e^{-n\int X_1 dx} \quad \text{or} \quad \frac{dv}{v^n} = X_2 e^{(1-n)\int X_1 dx}\,dx.$$

Hence $$v^{1-n} = (1-n)\int X_2 e^{(1-n)\int X_1 dx}\,dx, \quad \text{when} \quad n \neq 1,\text{*}$$

or $$y^{1-n} = (1-n)\,e^{(n-1)\int X_1 dx}\left[\int X_2 e^{(1-n)\int X_1 dx}\,dx\right]. \tag{6}$$

There is an analogous form for the second form of the equation.

The equation $(x^2y^3 + xy)\,dy = dx$ may be treated by this method by writing it as

$$\frac{dx}{dy} - yx = y^3x^2 \quad \text{so that} \quad Y_1 = -y,\ Y_2 = y^3,\ n = 2.$$

Then let $$x = ve^{-\int -y\,dy} = ve^{\frac{1}{2}y^2}.$$

Then $$\frac{dx}{dy} - yx = \frac{dv}{dy}\,e^{\frac{1}{2}y^2} + vye^{\frac{1}{2}y^2} - yve^{\frac{1}{2}y^2} = \frac{dv}{dy}\,e^{\frac{1}{2}y^2}$$

and $$\frac{dv}{dy}\,e^{\frac{1}{2}y^2} = y^3v^2e^{y^2} \quad \text{or} \quad \frac{dv}{v^2} = y^3e^{\frac{1}{2}y^2}\,dy,$$

and $$-\frac{1}{v} = (y^2-2)\,e^{\frac{1}{2}y^2} + C \quad \text{or} \quad \frac{1}{x} = 2 - y^2 + Ce^{-\frac{1}{2}y^2}.$$

This result could have been obtained by direct substitution in the formula

$$x^{1-n} = (1-n)\,e^{(n-1)\int Y_1 dy}\left[\int Y_2 e^{(1-n)\int Y_1 dy}\,dy\right],$$

but actually to carry the method through is far more instructive.

* If $n=1$, the variables are separated in the original equation.

EXERCISES

1. Solve the equations (variables immediately separable):

(α) $(1 + x)y + (1 - y)xy' = 0$, *Ans.* $xy = Ce^{x-y}$.

(β) $a(xdy + 2\,ydx) = xydy$, (γ) $\sqrt{1 - x^2}\,dy + \sqrt{1 - y^2}\,dx = 0$,

(δ) $(1 + y^2)\,dx - \left(y + \sqrt{1 + y}\right)(1 + x)^{\frac{3}{2}}\,dy = 0$.

2. By various ingenious changes of variable, solve:

(α) $(x + y)^2y' = a^2$, *Ans.* $x + y = a\tan(y/a + C)$.

(β) $(x - y^2)\,dx + 2\,xydy = 0$, ($\gamma$) $xdy - ydx = (x^2 + y^2)\,dx$,

(δ) $y' = x - y$, (ϵ) $yy' + y^2 + x + 1 = 0$.

3. Solve these homogeneous equations:

(α) $\left(2\sqrt{xy} - x\right)y' + y = 0$, *Ans.* $\sqrt{x/y} + \log y = C$.

(β) $xe^{\frac{y}{x}} + y - xy' = 0$, *Ans.* $y^2(x^2 + y^2) = Cx^6$.

(γ) $(x^2 + y^2)\,dy = xydx$, (δ) $xy' - y = \sqrt{x^2 + y^2}$.

4. Solve these Bernoulli or linear equations:

(α) $y' + y/x = y^2$, *Ans.* $xy\log Cx + 1 = 0$.

(β) $y' - y\csc x = \cos x - 1$, *Ans.* $y = \sin x + C\tan\frac{1}{2}x$.

(γ) $xy' + y = y^2\log x$, *Ans.* $y^{-1} = \log x + 1 + Cx$.

(δ) $(1 + y^2)\,dx + (\tan^{-1}y - x)\,dy = 0$ (ϵ) $ydx + (ax^2y^n - 2\,x)\,dy = 0$,

(ζ) $xy' - ay = x + 1$, (η) $yy' + \frac{1}{2}y^2 = \cos x$.

5. Show that the substitution $y = vx$ always separates the variables in the homogeneous equation $y' = \phi(y/x)$ and derive the general formula for the integral.

6. Let a differential equation be reducible to the form

$$\frac{dy}{dx} = \phi\left(\frac{a_1x + b_1y + c_1}{a_2x + b_2y + c_2}\right), \qquad \begin{array}{l} a_1b_2 - a_2b_1 \neq 0, \\ \text{or } a_1b_2 - a_2b_1 = 0. \end{array}$$

In case $a_1b_2 - a_2b_1 \neq 0$, the two lines $a_1x + b_1y + c_1 = 0$ and $a_2x + b_2y + c_2 = 0$ will meet in a point. Show that a transformation to this point as origin makes the new equation homogeneous and hence soluble. In case $a_1b_2 - a_2b_1 = 0$, the two lines are parallel and the substitution $z = a_2x + b_2y$ or $z = a_1x + b_1y$ will separate the variables.

7. By the method of Ex. 6 solve the equations:

(α) $(3y - 7x + 7)\,dx + (7y - 3x + 3)\,dy = 0$, *Ans.* $(y - x + 1)^2(y + x - 1)^7 = C$.

(β) $(2x + 3y - 5)\,y' + (3x + 2y - 5) = 0$, ($\gamma$) $(4x + 3y + 1)\,dx + (x + y + 1)\,dy = 0$,

(δ) $(2x + y) = y'(4x + 2y - 1)$, ($\epsilon$) $\dfrac{dy}{dx} = \left(\dfrac{x - y - 1}{2x - 2y + 1}\right)^2$.

8. Show that if the equation may be written as $yf(xy)\,dx + xg(xy)\,dy = 0$, where f and g are functions of the product xy, the substitution $v = xy$ will separate the variables.

9. By virtue of Ex. 8 integrate the equations:

(α) $(y + 2\,xy^2 - x^2y^3)\,dx + 2\,x^2ydy = 0$, *Ans.* $x + x^2y = C(1 - xy)$.

(β) $(y + xy^2)\,dx + (x - x^2y)\,dy = 0$, ($\gamma$) $(1 + xy)\,xy^2dx + (xy - 1)\,xdy = 0$.

10. By any method that is applicable solve the following. If more than one method is applicable, state what methods, and any apparent reasons for choosing one:

(α) $y' + y\cos x = y^n \sin 2x$, ($\beta$) $(2x^2y + 3y^3)\,dx = (x^3 + 2xy^2)\,dy$,
(γ) $(4x + 2y - 1)y' + 2x + y + 1 = 0$, ($\delta$) $yy' + xy^2 = x$,
(ϵ) $y'\sin y + \sin x \cos y = \sin x$, ($\zeta$) $\sqrt{a^2 + x^2}\,(1 - y') = x + y$,
(η) $(x^3y^3 + x^2y^2 + xy + 1)y + (x^3y^3 - x^2y^2 - xy + 1)xy'$, ($\theta$) $y' = \sin(x - y)$,
(ι) $xydy - y^2dx = (x + y)^2 e^{-\frac{y}{x}} dx$, ($\kappa$) $(1 - y^2)\,dx = axy(x + 1)\,dy$.

91. Integrating factors. If the equation $Mdx + Ndy = 0$ by a suitable rearrangement of the terms can be put in the form of a sum of total differentials of certain functions $u, v, \cdots$, say

$$du + dv + \cdots = 0, \quad \text{then} \quad u + v + \cdots = C \tag{7}$$

is surely the solution of the equation. In this case the equation is called an *exact differential equation.* It frequently happens that although the equation cannot itself be so arranged, yet the equation obtained from it by multiplying through with a certain factor $\mu(x, y)$ may be so arranged. The factor $\mu(x, y)$ is then called an *integrating factor* of the given equation. Thus in the case of variables separable, an integrating factor is $1/M_2N_1$; for

$$\frac{1}{M_2N_1}[M_1M_2\,dx + N_1N_2\,dy] = \frac{M_1(x)}{N_1(x)}\,dx + \frac{N_2(y)}{M_2(y)}\,dy = 0; \tag{8}$$

and the integration is immediate. Again, the linear equation may be treated by an integrating factor. Let

$$dy + X_1ydx = X_2dx \quad \text{and} \quad \mu = e^{\int X_1dx}; \tag{9}$$

then

$$e^{\int X_1dx}dy + X_1e^{\int X_1dx}ydx = e^{\int X_1dx}X_2dx \tag{10}$$

or

$$d\left[ye^{\int X_1dx}\right] = e^{\int X_1dx}X_2dx, \quad \text{and} \quad ye^{\int X_1dx} = \int e^{\int X_1dx}X_2dx. \tag{11}$$

In the case of variables separable the use of an integrating factor is therefore implied in the process of separating the variables. In the case of the linear equation the use of the integrating factor is somewhat shorter than the use of the substitution for separating the variables. In general it is not possible to hit upon an integrating factor by inspection and not practicable to obtain an integrating factor by analysis, but the integration of an equation is so simple when the factor is known, and the equations which arise in practice so frequently do have simple integrating factors, that it is worth while to examine the equation to see if the factor cannot be determined by inspection and trial. To aid in the work, the differentials of the simpler functions such as

$$dxy = xdy + ydx, \qquad \tfrac{1}{2}\, d(x^2 + y^2) = xdx + ydy,$$

$$d\frac{y}{x} = \frac{xdy - ydx}{x^2}, \qquad d\tan^{-1}\frac{x}{y} = \frac{ydx - xdy}{x^2 + y^2}, \tag{12}$$

should be borne in mind.

Consider the equation $(x^4e^x - 2mxy^2)\,dx + 2mx^2ydy = 0$. Here the first term x^4e^xdx will be a differential of a function of x no matter what function of x may be assumed as a trial μ. With $\mu = 1/x^4$ the equation takes the form

$$e^xdx + 2m\left(\frac{ydy}{x^2} - \frac{y^2dx}{x^3}\right) = de^x + md\frac{y^2}{x^2} = 0.$$

The integral is therefore seen to be $e^x + my^2/x^2 = C$ without more ado. It may be noticed that this equation is of the Bernoulli type and that an integration by that method would be considerably longer and more tedious than this use of an integrating factor.

Again, consider $(x + y)\,dx - (x - y)\,dy = 0$ and let it be written as

$$xdx + ydy + ydx - xdy = 0; \quad \text{try} \quad \mu = 1/(x^2 + y^2);$$

then $$\frac{xdx + ydy}{x^2 + y^2} + \frac{ydx - xdy}{x^2 + y^2} = 0 \quad \text{or} \quad \frac{1}{2}d\log(x^2 + y^2) + d\tan^{-1}\frac{x}{y} = 0,$$

and the integral is $\log\sqrt{x^2 + y^2} + \tan^{-1}(x/y) = C$. Here the terms $xdx + ydy$ strongly suggested $x^2 + y^2$ and the known form of the differential of $\tan^{-1}(x/y)$ corroborated the idea. This equation comes under the homogeneous type, but the use of the integrating factor considerably shortens the work of integration.

92. The attempt has been to write $Mdx + Ndy$ or $\mu(Mdx + Ndy)$ as the sum of total differentials $du + dv + \cdots$, that is, as the differential dF of the function $u + v + \cdots$, so that the solution of the equation $Mdx + Ndy = 0$ could be obtained as $F = C$. When the expressions are complicated, the attempt may fail in practice even where it theoretically should succeed. It is therefore of importance to establish conditions under which a differential expression like $Pdx + Qdy$ shall be the total differential dF of some function, and to find a means of obtaining F when the conditions are satisfied. This will now be done.

Suppose $$Pdx + Qdy = dF = \frac{\partial F}{\partial x}dx + \frac{\partial F}{\partial y}dy; \tag{13}$$

then $$P = \frac{\partial F}{\partial x}, \qquad Q = \frac{\partial F}{\partial y}, \qquad \frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x} = \frac{\partial^2 F}{\partial x \partial y}.$$

Hence if $Pdx + Qdy$ is a total differential dF, it follows (as in § 52) that the relation $P'_y = Q'_x$ must hold. Now conversely if this relation does hold, it may be shown that $Pdx + Qdy$ is the total differential of a function, and that this function is

$$F = \int_{x_0}^{x} P(x, y)\,dx + \int Q(x_0, y)\,dy$$

or

$$F = \int_{y_0}^{y} Q(x, y)\,dy + \int P(x, y_0)\,dx, \tag{14}$$

where the fixed value x_0 or y_0 will naturally be so chosen as to simplify the integrations as much as possible.

To show that these expressions may be taken as F it is merely necessary to compute their derivatives for identification with P and Q. Now

$$\frac{\partial F}{\partial x} = \frac{\partial}{\partial x}\int_{x_0}^{x} P(x, y)\,dx + \frac{\partial}{\partial x}\int Q(x_0, y)\,dy = P(x, y),$$

$$\frac{\partial F}{\partial y} = \frac{\partial}{\partial y}\int_{x_0}^{x} P(x, y)\,dx + \frac{\partial}{\partial y}\int Q(x_0, y)\,dy = \frac{\partial}{\partial y}\int Pdx + Q(x_0, y).$$

These differentiations, applied to the first form of F, require only the fact that the derivative of an integral is the integrand. The first turns out satisfactorily. The second must be simplified by interchanging the order of differentiation by y and integration by x (Leibniz's Rule, § 119) and by use of the fundamental hypothesis that $P'_y = Q'_x$.

$$\frac{\partial}{\partial y}\int_{x_0}^{x} Pdx + Q(x_0, y) = \int_{x_0}^{x} \frac{\partial P}{\partial y}\,dx + Q(x_0, y)$$

$$= \int_{x_0}^{x} \frac{\partial Q}{\partial x}\,dx + Q(x_0, y) = Q(x, y)\Big|_{x_0}^{x} + Q(x_0, y) = Q(x, y).$$

The identity of P and Q with the derivatives of F is therefore established. The second form of F would be treated similarly.

Show that $(x^2 + \log y)\,dx + x/y\,dy = 0$ is an exact differential equation and obtain the solution. Here it is first necessary to apply the test $P'_y = Q'_x$. Now

$$\frac{\partial}{\partial y}(x^2 + \log y) = \frac{1}{y} \quad \text{and} \quad \frac{\partial}{\partial x}\frac{x}{y} = \frac{1}{y}.$$

Hence the test is satisfied and the integral is obtained by applying the formula:

$$\int_0^x (x^2 + \log y)\,dx + \int \frac{0}{y}\,dy = \frac{1}{3}x^3 + x \log y = C$$

or

$$\int_1^y \frac{x}{y}\,dy + \int (x^2 + \log 1)\,dx = x \log y + \frac{1}{3}x^3 = C.$$

It should be noticed that the choice of $x_0 = 0$ simplifies the integration in the first case because the substitution of the lower limit 0 is easy and because the second integral vanishes. The choice of $y_0 = 1$ introduces corresponding simplifications in the second case.

Derive the *partial differential equation which any integrating factor of the differential equation $Mdx + Ndy = 0$ must satisfy*. If μ is an integrating factor, then

$$\mu Mdx + \mu Ndy = dF \quad \text{and} \quad \frac{\partial \mu M}{\partial y} = \frac{\partial \mu N}{\partial x}.$$

Hence

$$M\frac{\partial \mu}{\partial y} - N\frac{\partial \mu}{\partial x} = \mu\left(\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y}\right) \tag{15}$$

is the desired equation. To determine the integrating factor by solving this equation would in general be as difficult as solving the original equation; in some special cases, however, this equation is useful in determining μ.

93. It is now convenient to tabulate a list of different types of differential equations for which an integrating factor of a standard form can be given. With the knowledge of the factor, the equations may then be integrated by (14) or by inspection.

EQUATION $Mdx + Ndy = 0$:	FACTOR μ:
I. Homogeneous $Mdx + Ndy = 0$,	$\dfrac{1}{Mx + Ny}$.
II. Bernoulli $dy + X_1ydx = X_2y^ndx$,	$y^{-n}e^{(1-n)\int X_1dx}$.
III. $M = yf(xy)$, $N = xg(xy)$,	$\dfrac{1}{Mx - Ny}$.
IV. If $\dfrac{\frac{\partial M}{\partial y} - \frac{\partial N}{\partial x}}{N} = f(x)$,	$e^{\int f(x)dx}$.
V. If $\dfrac{\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y}}{M} = f(y)$,	$e^{\int f(y)dy}$.
VI. Type $x^\alpha y^\beta(mydx + nxdy) = 0$,	$x^{km-1-\alpha}y^{kn-1-\beta}$, k arbitrary.
VII. $x^\alpha y^\beta(mydx + nxdy) + x^\gamma y^\delta(pydx + qxdy) = 0$,	$x^{km-1-\alpha}y^{kn-1-\beta}$, k determined.

The use of the integrating factor often is simpler than the substitution $y = vx$ in the homogeneous equation. It is practically identical with the substitution in the Bernoulli type. In the third type it is often shorter than the substitution. The remaining types have had no substitution indicated for them. The proofs that the assigned forms of the factor are right are given in the examples below or are left as exercises.

To show that $\mu = (Mx + Ny)^{-1}$ is an integrating factor for the homogeneous case, it is possible simply to substitute in the equation (15), which μ must satisfy, and show that the equation actually holds by virtue of the fact that M and N are homogeneous of the same degree, — this fact being used to simplify the result by Euler's Formula (30) of § 53. But it is easier to proceed directly to show

$$\frac{\partial}{\partial y}\frac{M}{Mx+Ny}=\frac{\partial}{\partial x}\left(\frac{N}{Mx+Ny}\right)\quad\text{or}\quad\frac{\partial}{\partial y}\left(\frac{1}{x}\frac{1}{1+\phi}\right)=\frac{\partial}{\partial x}\left(\frac{1}{y}\frac{\phi}{1+\phi}\right),\quad\text{where}\quad\phi=\frac{Ny}{Mx}.$$

Owing to the homogeneity, ϕ is a function of y/x alone. Differentiate.

$$\frac{\partial}{\partial y}\left(\frac{1}{x}\frac{1}{1+\phi}\right)=-\frac{1}{x}\frac{\phi'}{(1+\phi)^2}\frac{1}{x}=\frac{1}{y}\frac{\phi'}{(1+\phi)^2}\cdot\frac{-y}{x^2}=\frac{\partial}{\partial x}\left(\frac{1}{y}\frac{\phi}{1+\phi}\right).$$

As this is an evident identity, the theorem is proved.

To find the condition that the integrating factor may be a function of x only and to find the factor when the condition is satisfied, the equation (15) which μ satisfies may be put in the more compact form by dividing by μ.

$$M\frac{1}{\mu}\frac{\partial\mu}{\partial y}-N\frac{1}{\mu}\frac{\partial\mu}{\partial x}=\frac{\partial N}{\partial x}-\frac{\partial M}{\partial y}\quad\text{or}\quad M\frac{\partial\log\mu}{\partial y}-N\frac{\partial\log\mu}{\partial x}=\frac{\partial N}{\partial x}-\frac{\partial M}{\partial y}.\qquad(15')$$

Now if μ (and hence $\log\mu$) is a function of x alone, the first term vanishes and

$$\frac{d\log\mu}{dx}=\frac{M'_y-N'_x}{N}=f(x)\quad\text{or}\quad\log\mu=\int f(x)\,dx.$$

This establishes the rule of type IV above and further shows that in no other case can μ be a function of x alone. The treatment of type V is clearly analogous.

Integrate the equation $x^4y(3\,ydx+2\,xdy)+x^2(4\,ydx+3\,xdy)=0$. This is of type VII; an integrating factor of the form $\mu=x^\rho y^\sigma$ will be assumed and the exponents ρ, σ will be determined so as to satisfy the condition that the equation be an exact differential. Here

$$P=\mu M=3\,x^{\rho+4}y^{\sigma+2}+4\,x^{\rho+2}y^{\sigma+1},\qquad Q=\mu N=2\,x^{\rho+5}y^{\sigma+1}+3\,x^{\rho+3}y^{\sigma}.$$

Then
$$P'_y=3(\sigma+2)x^{\rho+4}y^{\sigma+1}+4(\sigma+1)x^{\rho+2}y^{\sigma}$$
$$=2(\rho+5)x^{\rho+4}y^{\sigma+1}+3(\rho+3)x^{\rho+2}y^{\sigma}=Q'_x.$$

Hence if $\quad 3(\sigma+2)=2(\rho+5)\quad$ and $\quad 4(\sigma+1)=3(\rho+3)$,

the relation $P'_y=Q'_x$ will hold. This gives $\sigma=2$, $\rho=1$. Hence $\mu=xy^2$,

and
$$\int_0^x(3\,x^5y^4+4\,x^3y^3)\,dx+\int 0dy=\tfrac{1}{2}x^6y^4+x^4y^3=C$$

is the solution. The work might be shortened a trifle by dividing through in the first place by x^2. Moreover the integration can be performed at sight without the use of (14).

94. Several of the most important facts relative to integrating factors and solutions of $Mdx+Ndy=0$ will now be stated as theorems and the proofs will be indicated below.

1. If an integrating factor is known, the corresponding solution may be found; and conversely if the solution is known, the corresponding integrating factor may be found. Hence the existence of either implies the existence of the other.

2. If $F=C$ and $G=C$ are two solutions of the equation, either must be a function of the other, as $G=\Phi(F)$; and any function of either is

a solution. If μ and ν are two integrating factors of the equation, the ratio μ/ν is either constant or a solution of the equation; and the product of μ by any function of a solution, as $\mu\Phi(F)$, is an integrating factor of the equation.

3. The normal derivative dF/dn of a solution obtained from the factor μ is the product $\mu\sqrt{M^2+N^2}$ (see § 48).

It has already been seen that if an integrating factor μ is known, the corresponding solution $F=C$ may be found by (14). Now if the solution is known, the equation

$$dF = F'_x dx + F'_y dy = \mu(Mdx + Ndy) \quad \text{gives} \quad F'_x = \mu M, \ F'_y = \mu N;$$

and hence μ may be found from either of these equations as the quotient of a derivative of F by a coefficient of the differential equation. The statement 1 is therefore proved. It may be remarked that the discussion of approximate solutions to differential equations (§§ 86–88), combined with the theory of limits (beyond the scope of this text), affords a demonstration that any equation $Mdx + Ndy = 0$, where M and N satisfy certain restrictive conditions, has a solution; and hence it may be inferred that such an equation has an integrating factor.

If μ be eliminated from the relations $F'_x = \mu M$, $F'_y = \mu N$ found above, it is seen that

$$MF'_y - NF'_x = 0, \quad \text{and similarly,} \quad MG'_y - NG'_x = 0, \tag{16}$$

are the conditions that F and G should be solutions of the differential equation. Now these are two simultaneous homogeneous equations of the first degree in M and N. If M and N are eliminated from them, there results the equation

$$F'_y G'_x - F'_x G'_y = 0 \quad \text{or} \quad \begin{vmatrix} F'_x & F'_y \\ G'_x & G'_y \end{vmatrix} = J(F, G) = 0, \tag{16'}$$

which shows (§ 62) that F and G are functionally related as required. To show that any function $\Phi(F)$ is a solution, consider the equation

$$M\Phi'_y - N\Phi'_x = (MF'_y - NF'_x)\,\Phi'.$$

As F is a solution, the expression $MF'_y - NF'_x$ vanishes by (16), and hence $M\Phi'_y - N\Phi'_x$ also vanishes, and Φ is a solution of the equation as is desired. The first half of 2 is proved.

Next, if μ and ν are two integrating factors, equation (15') gives

$$M\frac{\partial \log \mu}{\partial y} - N\frac{\partial \log \mu}{\partial x} = M\frac{\partial \log \nu}{\partial y} - N\frac{\partial \log \nu}{\partial x} \quad \text{or} \quad M\frac{\partial \log \mu/\nu}{\partial y} - N\frac{\partial \log \mu/\nu}{\partial x} = 0.$$

On comparing with (16) it then appears that $\log(\mu/\nu)$ must be a solution of the equation and hence μ/ν itself must be a solution. The inference, however, would not hold if μ/ν reduced to a constant. Finally if μ is an integrating factor leading to the solution $F=C$, then

$$dF = \mu(Mdx + Ndy), \quad \text{and hence} \quad \mu\Phi(F)(Mdx + Ndy) = d\int \Phi(F)\,dF.$$

It therefore appears that the factor $\mu\Phi(F)$ makes the equation an exact differential and must be an integrating factor. Statement 2 is therefore wholly proved.

The third proposition is proved simply by differentiation and substitution. For

$$\frac{dF}{dn} = \frac{\partial F}{\partial x}\frac{dx}{dn} + \frac{\partial F}{\partial y}\frac{dy}{dn} = \mu M \frac{dx}{dn} + \mu N \frac{dy}{dn}.$$

And if τ denotes the inclination of the curve $F = C$, it follows that

$$\tan\tau = \frac{dy}{dx} = -\frac{M}{N}, \qquad \sin\tau = \frac{dy}{dn} = \frac{N}{\sqrt{M^2+N^2}}, \qquad -\cos\tau = \frac{dx}{dn} = \frac{M}{\sqrt{M^2+N^2}}.$$

Hence $dF/dn = \mu\sqrt{M^2+N^2}$ and the proposition is proved.

EXERCISES

1. Find the integrating factor by inspection and integrate:

(α) $xdy - ydx = (x^2+y^2)dx$, (β) $(y^2 - xy)dx + x^2dy = 0$,
(γ) $ydx - xdy + \log x dx = 0$, ($\delta$) $y(2xy + e^x)dx - e^xdy = 0$,
(ϵ) $(1+xy)ydx + (1-xy)xdy = 0$, ($\zeta$) $(x - y^2)dx + 2xydy = 0$,
(η) $(xy^2 + y)dx - xdy = 0$, (θ) $a(xdy + 2ydx) = xydy$,
(ι) $(x^2+y^2)(xdx + ydy) + \sqrt{1+(x^2+y^2)}(ydx - xdy) = 0$,
(κ) $x^2ydx - (x^3+y^3)dy = 0$, (λ) $xdy - ydx = x\sqrt{x^2-y^2}dy$.

2. Integrate these linear equations with an integrating factor:

(α) $y' + ay = \sin bx$, (β) $y' + y\cot x = \sec x$,
(γ) $(x+1)y' - 2y = (x+1)^4$, (δ) $(1+x^2)y' + y = e^{\tan^{-1}x}$,

and (β), (δ), (ζ) of Ex. 4, p. 206.

3. Show that the expression given under II, p. 210, is an integrating factor for the Bernoulli equation, and integrate the following equations by that method:

(α) $y' - y\tan x = y^4\sec x$, (β) $3y^2y' + y^3 = x - 1$,
(γ) $y' + y\cos x = y^n \sin 2x$, ($\delta$) $dx + 2xydy = 2ax^3y^3dy$,

and (α), (γ), (ϵ), (η) of Ex. 4, p. 206.

4. Show the following are exact differential equations and integrate:

(α) $(3x^2+6xy^2)dx + (6x^2y + 4y^2)dy = 0$, ($\beta$) $\sin x\cos ydx + \cos x\sin ydy = 0$,
(γ) $(6x - 2y + 1) + (2y - 2x - 3)dy = 0$, ($\delta$) $(x^3 + 3xy^2)dx + (y^3 + 3x^2y)dy = 0$,
(ϵ) $\dfrac{2xy+1}{y}dx + \dfrac{y-x}{y^2}dy = 0$, ($\zeta$) $\left(1 + e^{\frac{x}{y}}\right)dx + e^{\frac{x}{y}}\left(1 - \dfrac{x}{y}\right)dy = 0$,
(η) $e^x(x^2+y^2+2x)dx + 2ye^xdy = 0$, ($\theta$) $(y\sin x - 1)dx + (y - \cos x)dy = 0$.

5. Show that $(Mx - Ny)^{-1}$ is an integrating factor for type III. Determine the integrating factors of the following equations, thus render them exact, and integrate:

(α) $(y+x)dx + xdy = 0$, (β) $(y^2 - xy)dx + x^2dy = 0$,
(γ) $(x^2+y^2)dx - 2xydy = 0$, (δ) $(x^2y^2 + xy)ydx + (x^2y^2 - 1)xdy = 0$,
(ϵ) $(\sqrt{xy} - 1)xdy - (\sqrt{xy} + 1)ydx = 0$, ($\zeta$) $x^3dx + (3x^2y + 2y^3)dy = 0$,

and Exs. 3 and 9, p. 206.

6. Show that the factor given for type VI is right, and that the form given for type VII is right if k satisfies $k(qm - pn) = q(\alpha - \gamma) - p(\beta - \delta)$.

7. Integrate the following equations of types IV–VII:

(α) $(y^4 + 2y)dx + (xy^3 + 2y^4 - 4x)dy = 0$, ($\beta$) $(x^2 + y^2 + 1)dx - 2xydy = 0$,
(γ) $(3x^2 + 6xy\,3y^2)dx + (2x^2 + 3xy)dy = 0$, ($\delta$) $(2x^2y^2 + y) - (x^3y - 3x)y' = 0$,
(ϵ) $(2x^2y - 3y^4)dx + (3x^3 + 2xy^3)dy = 0$,
(ζ) $(2 - y')\sin(3x - 2y) + y'\sin(x - 2y) = 0$.

8. By virtue of proposition 2 above, it follows that if an equation is exact and homogeneous, or exact and has the variables separable, or homogeneous and under types IV–VII, so that two different integrating factors may be obtained, the solution of the equation may be obtained without integration. Apply this to finding the solutions of Ex. 4 (β), (δ), (γ); Ex. 5 (α), (γ).

9. Discuss the apparent exceptions to the rules for types I, III, VII, that is, when $Mx + Ny = 0$ or $Mx - Ny = 0$ or $qm - pn = 0$.

10. Consider this rule for integrating $Mdx + Ndy = 0$ when the equation is known to be exact: Integrate Mdx regarding y as constant, differentiate the result regarding y as variable, and subtract from N; then integrate the difference with respect to y. In symbols,

$$C = \int (Mdx + Ndy) = \int Mdx + \int \left(N - \frac{\partial}{\partial y}\int Mdx\right)dy.$$

Apply this instead of (14) to Ex. 4. Observe that in no case should either this formula or (14) be applied when the integral is obtainable by inspection.

95. Linear equations with constant coefficients. The type

$$a_0\frac{d^ny}{dx^n} + a_1\frac{d^{n-1}y}{dx^{n-1}} + \cdots + a_{n-1}\frac{dy}{dx} + a_ny = X(x) \tag{17}$$

of differential equation of the nth order which is of the first degree in y and its derivatives is called a *linear* equation. For the present only the case where the coefficients $a_0, a_1, \cdots, a_{n-1}, a_n$ are constant will be treated, and for convenience it will be assumed that the equation has been divided through by a_0 so that the coefficient of the highest derivative is 1. Then if differentiation be denoted by D, the equation may be written *symbolically* as

$$(D^n + a_1D^{n-1} + \cdots + a_{n-1}D + a_n)y = X, \tag{17'}$$

where the symbol D combined with constants follows many of the laws of ordinary algebraic quantities (see § 70).

The simplest equation would be of the first order. Here

$$\frac{dy}{dx} - a_1y = X \quad \text{and} \quad y = e^{a_1x}\int e^{-a_1x}Xdx, \tag{18}$$

as may be seen by reference to (11) or (6). Now if $D - a_1$ be treated as an algebraic symbol, the solution may be indicated as

$$(D - a_1)y = X \quad \text{and} \quad y = \frac{1}{D - a_1}X, \tag{18'}$$

where the operator $(D - a_1)^{-1}$ is the *inverse* of $D - a_1$. The solution which has just been obtained shows that the interpretation which must be assigned to the inverse operator is

$$\frac{1}{D - a_1}(*) = e^{a_1 x}\int e^{-a_1 x}(*)\,dx, \tag{19}$$

where $(*)$ denotes the function of x upon which it operates. That the integrating operator is the inverse of $D - a_1$ may be proved by direct differentiation (see Ex. 7, p. 152).

This operational method may at once be extended to obtain the solution of equations of higher order. For consider

$$\frac{d^2y}{dx^2} + a_1\frac{dy}{dx} + a_2 y = X \quad \text{or} \quad (D^2 + a_1 D + a_2)\, y = X. \tag{20}$$

Let α_1 and α_2 be the roots of the equation $D^2 + a_1 D + a_2 = 0$ so that the differential equation may be written in the form

$$[D^2 - (\alpha_1 + \alpha_2)\, D + \alpha_1\alpha_2]\, y = X \quad \text{or} \quad (D - \alpha_1)(D - \alpha_2)\, y = X. \tag{20'}$$

The solution may now be evaluated by a succession of steps as

$$(D - \alpha_2)\, y = \frac{1}{D - \alpha_1} X = e^{\alpha_1 x}\int e^{-\alpha_1 x} X dx,$$

$$y = \frac{1}{D - \alpha_2}\left[\frac{1}{D - \alpha_1} X\right] = e^{\alpha_2 x}\int e^{-\alpha_2 x}\left[e^{\alpha_1 x}\int e^{-\alpha_1 x} X dx\right]$$

or

$$y = e^{\alpha_2 x}\int e^{(\alpha_1 - \alpha_2)x}\left[\int e^{-\alpha_1 x} X dx\right] dx. \tag{20''}$$

The solution of the equation is thus reduced to quadratures.

The extension of the method to an equation of any order is immediate. The first step in the solution is to solve the equation

$$D^n + a_1 D^{n-1} + \cdots + a_{n-1} D + a_n = 0$$

so that the differential equation may be written in the form

$$(D - \alpha_1)(D - \alpha_2)\cdots(D - \alpha_{n-1})(D - \alpha_n)\, y = X; \tag{17''}$$

whereupon the solution is comprised in the formula

$$y = e^{\alpha_n x}\int e^{(\alpha_{n-1} - \alpha_n)x}\int \cdots \int e^{(\alpha_1 - \alpha_2)x}\int e^{-\alpha_1 x} X\,(dx)^n, \tag{17'''}$$

where the successive integrations are to be performed by beginning upon the extreme right and working toward the left. Moreover, it appears that if the operators $D - \alpha_n$, $D - \alpha_{n-1}, \cdots, D - \alpha_2$, $D - \alpha_1$ were successively applied to this value of y, they would undo the work here done and lead back to the original equation. As n integrations are required, there will occur n arbitrary constants of integration in the answer for y.

As an example consider the equation $(D^3 - 4D)y = x^2$. Here the roots of the algebraic equation $D^3 - 4D = 0$ are $0, 2, -2$, and the solution for y is

$$y = \frac{1}{D}\frac{1}{D-2}\frac{1}{D+2}x^2 = \int e^{2x}\int e^{-2x}e^{-2x}\int e^{2x}x^2(dx)^3.$$

The successive integrations are very simple by means of a table. Then

$$\int e^{2x}x^2dx = \tfrac{1}{2}x^2e^{2x} - \tfrac{1}{2}xe^{2x} + \tfrac{1}{4}e^{2x} + C_1,$$

$$\int e^{-4x}\int e^{2x}x^2(dx)^2 = \int(\tfrac{1}{2}x^2e^{-2x} - \tfrac{1}{2}xe^{-2x} + \tfrac{1}{4}e^{-2x} + C_1e^{-4x})\,dx$$
$$= -\tfrac{1}{4}x^2e^{-2x} - \tfrac{1}{8}e^{-2x} + C_1e^{-4x} + C_2,$$

$$y = \int e^{2x}\int e^{-4x}\int e^{2x}x^2(dx)^3 = \int(-\tfrac{1}{4}x^2 - \tfrac{1}{8} + C_1e^{-2x} + C_2e^{2x})\,dx$$
$$= -\tfrac{1}{12}x^3 - \tfrac{1}{8}x + C_1e^{-2x} + C_2e^{2x} + C_3.$$

This is the solution. It may be noted that in integrating a term like C_1e^{-4x} the result may be written as C_1e^{-4x}, for the reason that C_1 is arbitrary anyhow; and, moreover, if the integration had introduced any terms such as $2e^{-2x}$, $\frac{1}{2}e^{2x}$, 5, these could be combined with the terms C_1e^{-2x}, C_2e^{2x}, C_3 to simplify the form of the results.

In case the roots are imaginary the procedure is the same. Consider

$$\frac{d^2y}{dx^2} + y = \sin x \quad\text{or}\quad (D^2+1)y = \sin x \quad\text{or}\quad (D+i)(D-i)y = \sin x.$$

Then $$y = \frac{1}{D-i}\frac{1}{D+i}\sin x = e^{ix}\int e^{-2ix}\int e^{ix}\sin x\,(dx)^2, \qquad i = \sqrt{-1}.$$

The formula for $\int e^{ax}\sin bx dx$, as given in the tables, is not applicable when $a^2 + b^2 = 0$, as is the case here, because the denominator vanishes. It therefore becomes expedient to write $\sin x$ in terms of exponentials. Then

$$y = e^{ix}\int e^{-2ix}\int e^{ix}\frac{e^{ix}-e^{-ix}}{2i}(dx)^2; \quad\text{for}\quad \sin x = \frac{e^{ix}-e^{-ix}}{2i}.$$

Now $$\frac{1}{2i}e^{ix}\int e^{-2ix}\int(e^{2ix}-1)(dx)^2 = \frac{1}{2i}e^{ix}\int e^{-2ix}\left[\frac{1}{2i}e^{2ix} - x + C_1\right]dx$$
$$= \frac{1}{2i}e^{ix}\left[\frac{1}{2i}x + \frac{1}{2i}e^{-2ix}x - \frac{1}{4}e^{-2ix} + C_1e^{-2ix} + C_2\right]$$
$$= -\frac{x}{2}\frac{e^{ix}+e^{-ix}}{2} + C_1e^{-ix} + C_2e^{ix}.$$

Now $$C_1e^{-ix} + C_2e^{ix} = (C_2 + C_1)\frac{e^{ix}+e^{-ix}}{2} + (C_2 - C_1)\,i\,\frac{e^{ix}-e^{-ix}}{2i}.$$

Hence this expression may be written as $C_1\cos x + C_2\sin x$, and then

$$y = -\tfrac{1}{2}x\cos x + C_1\cos x + C_2\sin x.$$

The solution of such equations as these gives excellent opportunity to cultivate the art of manipulating trigonometric functions through exponentials (§ 74).

96. The general method of solution given above may be considerably simplified in case the function $X(x)$ has certain special forms. In the first place suppose $X = 0$, and let the equation be $P(D)y = 0$, where $P(D)$ denotes the symbolic polynomial of the nth degree in D. Suppose the roots of $P(D) = 0$ are $\alpha_1, \alpha_2, \cdots, \alpha_k$ and their respective multiplicities are $m_1, m_2, \cdots, m_k$, so that

$$(D - \alpha_k)^{m_k} \cdots (D - \alpha_2)^{m_2}(D - \alpha_1)^{m_1} y = 0$$

is the form of the differential equation. Now, as above, if

$$(D - \alpha_1)^{m_1} y = 0, \quad \text{then} \quad y = \frac{1}{(D - \alpha_1)^{m_1}} 0 = e^{\alpha_1 x} \int \cdots \int 0 (dx)^{m_1}.$$

Hence $\qquad y = e^{\alpha_1 x}(C_1 + C_2 x + C_3 x^2 + \cdots + C_{m_1} x^{m_1 - 1})$

is annihilated by the application of the operator $(D - \alpha_1)^{m_1}$, and therefore by the application of the whole operator $P(D)$, and must be a solution of the equation. As the factors in $P(D)$ may be written so that any one of them, as $(D - \alpha_i)^{m_i}$, comes last, it follows that to each factor $(D - \alpha_i)^{m_i}$ will correspond a solution

$$y_i = e^{\alpha_i x}(C_{i1} + C_{i2} x + \cdots + C_{im_i} x^{m_i - 1}), \qquad P(D) y_i = 0,$$

of the equation. Moreover the sum of all these solutions,

$$y = \sum_{i=1}^{i=k} e^{\alpha_i x}(C_{i1} + C_{i2} x + \cdots + C_{im_i} x^{m_i - 1}), \tag{21}$$

will be a solution of the equation; for in applying $P(D)$ to y,

$$P(D) y = P(D) y_1 + P(D) y_2 + \cdots + P(D) y_k = 0.$$

Hence the general rule may be stated that: *The solution of the differential equation $P(D)y = 0$ of the nth order may be found by multiplying each $e^{\alpha x}$ by a polynomial of the mth degree in x (where α is a root of the equation $P(D) = 0$ of multiplicity m and where the coefficients of the polynomial are arbitrary) and adding the results.* Two observations may be made. First, the solution thus found contains n arbitrary constants and may therefore be considered as the general solution; and second, if there are imaginary roots for $P(D) = 0$, *the exponentials arising from the pure imaginary parts of the roots may be converted into trigonometric functions.*

As an example take $(D^4 + 2D^3 + D^2) y = 0$. The roots are 1, 1, 0, 0. Hence the solution is

$$y = e^x(C_1 + C_2 x) + (C_3 + C_4 x).$$

Again if $(D^4 + 4) y = 0$, the roots of $D^4 + 4 = 0$ are $\pm 1 \pm i$ and the solution is

$$y = C_1 e^{(1+i)x} + C_2 e^{(1-i)x} + C_3 e^{(-1+i)x} + C_4 e^{(-1-i)x}$$

or
$$y = e^x(C_1e^{ix} + C_2e^{-ix}) + e^{-x}(C_3e^{ix} + C_4e^{-ix})$$
$$= e^x(C_1 \cos x + C_2 \sin x) + e^{-x}(C_3 \cos x + C_4 \sin x),$$

where the new C's are not identical with the old C's. Another form is

$$y = e^x A \cos(x + \gamma) + e^{-x} B \cos(x + \delta),$$

where γ and δ, A and B, are arbitrary constants. For

$$C_1 \cos x + C_2 \sin x = \sqrt{C_1^2 + C_2^2}\left[\frac{C_1}{\sqrt{C_1^2 + C_2^2}}\cos x + \frac{C_2}{\sqrt{C_1^2 + C_2^2}}\sin x\right],$$

and if $\gamma = \tan^{-1}\left(-\frac{C_2}{C_1}\right)$, then $C_1 \cos x + C_2 \sin x = \sqrt{C_1^2 + C_2^2}\cos(x + \gamma)$.

Next if x is not zero but *if any one solution I can be found so that $P(D)I = X$, then a solution containing n arbitrary constants may be found by adding to I the solution of $P(D)y = 0$.* For if

$$P(D)I = X \quad \text{and} \quad P(D)y = 0, \quad \text{then} \quad P(D)(I + y) = X.$$

It therefore remains to devise means for finding one solution I. This solution I may be found by the long method of (17‴), where the integration may be shortened by omitting the constants of integration since only one, and not the general, value of the solution is needed. In the most important cases which arise in practice there are, however, some very short cuts to the solution I. The solution I of $P(D)y = X$ is called the *particular integral* of the equation and the general solution of $P(D)y = 0$ is called the *complementary function* for the equation $P(D)y = X$.

Suppose that X *is a polynomial in* x. Solve symbolically, arrange $P(D)$ in ascending powers of D, and divide out to powers of D equal to the order of the polynomial X. Then

$$P(D)I = X, \qquad I = \frac{1}{P(D)}X = \left[Q(D) + \frac{R(D)}{P(D)}\right]X, \tag{22}$$

where the remainder $R(D)$ is of *higher* order in D than X in x. Then

$$P(D)I = P(D)Q(D)X + R(D)X, \qquad R(D)X = 0.$$

Hence $Q(D)X$ may be taken as I, since $P(D)Q(D)X = P(D)I = X$. By this method the solution I may be found, when X is a polynomial, *as rapidly as $P(D)$ can be divided into* 1; the solution of $P(D)y = 0$ may be written down by (21); and the sum of I and this will be the required solution of $P(D)y = X$ containing n constants.

As an example consider $(D^3 + 4D^2 + 3D)y = x^2$. The work is as follows:

$$I = \frac{1}{3D + 4D^2 + D^3}x^2 = \frac{1}{D}\frac{1}{3 + 4D + D^2}x^2 = \frac{1}{D}\left[\frac{1}{3} - \frac{4}{9}D + \frac{13}{27}D^2 + \frac{R(D)}{P(D)}\right]x^2.$$

Hence $$I = Q(D)x^2 = \frac{1}{D}\left(\frac{1}{3} - \frac{4}{9}D + \frac{13}{27}D^2\right)x^2 = \frac{1}{9}x^3 - \frac{4}{9}x^2 + \frac{26}{27}x.$$

For $D^3 + 4D^2 + 3D = 0$ the roots are $0, -1, -3$ and the complementary function or solution of $P(D)y = 0$ would be $C_1 + C_2e^{-x} + C_3e^{-3x}$. Hence the solution of the equation $P(D)y = x^2$ is

$$y = C_1 + C_2e^{-x} + C_3e^{-3x} + \tfrac{1}{9}x^3 - \tfrac{4}{9}x^2 + \tfrac{26}{27}x.$$

It should be noted that in this example D is a factor of $P(D)$ and has been taken out before dividing; this shortens the work. Furthermore note that, in interpreting $1/D$ as integration, the constant may be omitted because any one value of I will do.

97. Next suppose that $X = Ce^{ax}$. Now $De^{ax} = ae^{ax}$, $D^ke^{ax} = a^ke^{ax}$,

and $$P(D)e^{ax} = P(a)e^{ax}; \quad \text{hence} \quad P(D)\left[\frac{C}{P(a)}e^{ax}\right] = Ce^{ax}.$$

But $$P(D)I = Ce^{ax}, \quad \text{and hence} \quad I = \frac{C}{P(a)}e^{ax} \tag{23}$$

is clearly a solution of the equation, provided a is not a root of $P(D) = 0$. If $P(a) = 0$, the division by $P(a)$ is impossible and the quest for I has to be directed more carefully. Let a be a root of multiplicity m so that $P(D) = (D - a)^mP_1(D)$. Then

$$P_1(D)(D - a)^mI = Ce^{ax}, \qquad (D - a)^mI = \frac{C}{P_1(a)}e^{ax},$$

and $$I = \frac{C}{P_1(a)}e^{ax}\int\cdots\int(dx)^m = \frac{Ce^{ax}x^m}{P_1(a)\,m!}. \tag{23'}$$

For in the integration the constants may be omitted. It follows that when $X = Ce^{ax}$, the solution I may be found *by direct substitution.*

Now if X broke up into the sum of terms $X = X_1 + X_2 + \cdots$ and if solutions $I_1, I_2, \cdots$ were determined for each of the equations $P(D)I_1 = X_1$, $P(D)I_2 = X_2, \cdots$, the solution I corresponding to X would be the sum $I_1 + I_2 + \cdots$. Thus it is seen that the above short methods apply to equations in which X is a sum of terms of the form Cx^m or Ce^{ax}.

As an example consider $(D^4 - 2D^2 + 1)y = e^x$. The roots are $1, 1, -1, -1$, and $a = 1$. Hence the solution for I is written as

$$(D + 1)^2(D - 1)^2I = e^x, \qquad (D - 1)^2I = \tfrac{1}{4}e^x, \qquad I = \tfrac{1}{8}e^xx^2.$$

Then $$y = e^x(C_1 + C_2x) + e^{-x}(C_3 + C_4x) + \tfrac{1}{8}e^xx^2.$$

Again consider $(D^2 - 5D + 6)y = x + e^{mx}$. To find the I_1 corresponding to x, divide.

$$I_1 = \frac{1}{6 - 5D + D^2}x = \left(\frac{1}{6} + \frac{5}{36}D + \cdots\right)x = \frac{1}{6}x + \frac{5}{36}.$$

To find the I_2 corresponding to e^{mx}, substitute. There are three cases,

$$I_2 = \frac{1}{m^2 - 5m + 6}e^{mx}, \qquad I_2 = xe^{3x}, \qquad I_2 = -xe^{2x},$$

according as m is neither 2 nor 3, or is 3, or is 2. Hence for the complete solution,

$$y = C_1e^{3x} + C_2e^{2x} + \frac{1}{6}x + \frac{5}{36} + \frac{1}{m^2 - 5m + 6}e^{mx},$$

when m is neither 2 nor 3; but in these special cases the results are

$$y = C_1e^{3x} + C_2e^{2x} + \tfrac{1}{6}x + \tfrac{5}{36} - xe^{2x}, \qquad y = C_1e^{3x} + C_2e^{2x} + \tfrac{1}{6}x + \tfrac{5}{36} + xe^{3x}.$$

The next case to consider is where X is of the form $\cos\beta x$ or $\sin\beta x$. If these trigonometric functions be expressed in terms of exponentials, the solution may be conducted by the method above; and this is perhaps the best method when $\pm\beta i$ are roots of the equation $P(D) = 0$. It may be noted that this method would apply also to the case where X might be of the form $e^{ax}\cos\beta x$ or $e^{ax}\sin\beta x$. Instead of splitting the trigonometric functions into two exponentials, it is possible to combine two trigonometric functions into an exponential. Thus, consider the equations

$$P(D)y = e^{ax}\cos\beta x, \qquad P(D)y = e^{ax}\sin\beta x,$$

and
$$P(D)y = e^{ax}(\cos\beta x + i\sin\beta x) = e^{(a+\beta i)x}. \tag{24}$$

The solution I of this last equation may be found and split into its real and imaginary parts, of which the real part is the solution of the equation involving the cosine, and the imaginary part the sine.

When X has the form $\cos\beta x$ or $\sin\beta x$ and $\pm\beta i$ are not roots of the equation $P(D) = 0$, there is a very short method of finding I. For

$$D^2\cos\beta x = -\beta^2\cos\beta x \quad \text{and} \quad D^2\sin\beta x = -\beta^2\sin\beta x.$$

Hence if $P(D)$ be written as $P_1(D^2) + DP_2(D^2)$ by collecting the even terms and the odd terms so that P_1 and P_2 are both even in D, the solution may be carried out symbolically as

$$I = \frac{1}{P(D)}\cos x = \frac{1}{P_1(D^2) + DP_2(D^2)}\cos x = \frac{1}{P_1(-\beta^2) + DP_2(-\beta^2)}\cos x,$$

or
$$I = \frac{P_1(-\beta^2) - DP_2(-\beta^2)}{[P_1(-\beta^2)]^2 + \beta^2[P_2(-\beta^2)]^2}\cos x. \tag{25}$$

By this device of substitution and of rationalization as if D were a surd, the differentiation is transferred to the numerator and can be performed. This method of procedure may be justified directly, or it may be made to depend upon that of the paragraph above.

Consider the example $(D^2 + 1)y = \cos x$. Here $\beta i = i$ is a root of $D^2 + 1 = 0$. As an operator D^2 is equivalent to -1, and the rationalization method will not work. If the first solution be followed, the method of solution is

$$I = \frac{1}{D^2+1}\frac{e^{ix}}{2} + \frac{1}{D^2+1}\frac{e^{-ix}}{2} = \frac{1}{D-i}\frac{e^{ix}}{4i} - \frac{1}{D+i}\frac{e^{-ix}}{4i} = \frac{1}{4i}[xe^{ix} - xe^{-ix}] = \frac{1}{2}x\sin x.$$

If the second suggestion be followed, the solution may be found as follows:

$$(D^2 + 1) I = \cos x + i \sin x = e^{ix}, \qquad I = \frac{1}{D^2 + 1} e^{ix} = \frac{xe^{ix}}{2i}.$$

Now $$I = \frac{x}{2i}(\cos x + i \sin x) = \frac{1}{2} x \sin x - \frac{1}{2} ix \cos x.$$

Hence $I = \frac{1}{2} x \sin x$ for $(D^2 + 1) I = \cos x,$

and $I = -\frac{1}{2} x \cos x$ for $(D^2 + 1) I = \sin x.$

The complete solution is $y = C_1 \cos x + C_2 \sin x + \frac{1}{2} x \sin x,$

and for $(D^2 + 1) y = \sin x,$ $y = C_1 \cos x + C_2 \sin x - \frac{1}{2} x \cos x.$

As another example take $(D^2 - 3D + 2) y = \cos x$. The roots are 1, 2, neither is equal to $\pm \beta i = \pm i$, and the method of rationalization is practicable. Then

$$I = \frac{1}{D^2 - 3D + 2} \cos x = \frac{1}{1 - 3D} \cos x = \frac{1 + 3D}{10} \cos x = \frac{1}{10}(\cos x - 3 \sin x).$$

The complete solution is $y = C_1 e^{-x} + C_2 e^{-2x} + \frac{1}{10}(\cos x - 3 \sin x)$. The extreme simplicity of this substitution-rationalization method is noteworthy.

EXERCISES

1. By the general method solve the equations:

(α) $\frac{d^2y}{dx^2} + 4\frac{dy}{dx} + 3y = 2e^{2x},$ (β) $\frac{d^3y}{dx^3} - 3\frac{d^2y}{dx^2} + 3\frac{dy}{dx} - y = e^x,$

(γ) $(D^2 - 4D + 2) y = x,$ (δ) $(D^3 + D^2 - 4D + 4) y = x,$

(ϵ) $(D^3 + 5D^2 + 6D) y = x,$ (ζ) $(D^2 + D + 1) y = xe^x,$

(η) $(D^2 + D + 1) y = \sin 2x,$ (θ) $(D^2 - 4) y = x + e^{2x},$

(ι) $(D^2 + 3D + 2) y = x + \cos x,$ (κ) $(D^4 - 4D^2) y = 1 - \sin x,$

(λ) $(D^2 + 1) y = \cos x,$ (μ) $(D^2 + 1) y = \sec x,$ (ν) $(D^2 + 1) y = \tan x.$

2. By the rule write the solutions of these equations:

(α) $(D^2 + 3D + 2) y = 0,$ (β) $(D^3 + 3D^2 + D - 5) y = 0,$

(γ) $(D - 1)^3 y = 0,$ (δ) $(D^4 + 2D^2 + 1) y = 0,$

(ϵ) $(D^3 - 3D^2 + 4) y = 0,$ (ζ) $(D^4 - D^3 - 9D^2 - 11D - 4) y = 0,$

(η) $(D^3 - 6D^2 + 9D) y = 0,$ (θ) $(D^4 - 4D^3 + 8D^2 - 8D + 4) y = 0,$

(ι) $(D^5 - 2D^4 + D^3) y = 0,$ (κ) $(D^3 - D^2 + D) y = 0,$

(λ) $(D^4 - 1)^2 y = 0,$ (μ) $(D^5 - 13D^3 + 26D^2 + 82D + 104) y = 0.$

3. By the short method solve (γ), (δ), (ϵ) of Ex. 1, and also:

(α) $(D^4 - 1) y = x^4,$ (β) $(D^3 - 6D^2 + 11D - 6) y = x,$

(γ) $(D^3 + 3D^2 + 2D) y = x^2,$ (δ) $(D^3 - 3D^2 - 6D + 8) y = x,$

(ϵ) $(D^3 + 8) y = x^4 + 2x + 1,$ (ζ) $(D^3 - 3D^2 - D + 3) y = x^2,$

(η) $(D^4 - 2D^3 + D^2) y = x,$ (θ) $(D^4 + 2D^3 + 3D^2 + 2D + 1) y = 1 + x + x^2,$

(ι) $(D^3 - 1) y = x^2,$ (κ) $(D^4 - 2D^3 + D^2) y = x^3.$

4. By the short method solve (α), (β), (θ) of Ex. 1, and also:

(α) $(D^2 - 3D + 2) y = e^x,$ (β) $(D^4 - D^3 - 3D^2 + 5D - 2) y = e^{3x},$

(γ) $(D^2 - 2D + 1) y = e^x,$ (δ) $(D^3 - 3D^2 + 4) y = e^{3x},$

(ϵ) $(D^2 + 1) y = 2e^x + x^3 - x,$ (ζ) $(D^3 + 1) y = 3 + e^{-x} + 5e^{2x},$

(η) $(D^4 + 2D^2 + 1) y = e^x + 4,$ (θ) $(D^3 + 3D^2 + 3D + 1) y = 2e^{-x},$

(ι) $(D^2 - 2D) y = e^{2x} + 1,$ (κ) $(D^3 + 2D^2 + D) y = e^{2x} + x^2 + x,$

(λ) $(D^2 - a^2) y = e^{ax} + e^{bx},$ (μ) $(D^2 - 2aD + a^2) y = e^x + 1.$

5. Solve by the short method (η), (ι), (κ) of Ex. 1, and also:

(α) $(D^2 - D - 2)y = \sin x$,
(β) $(D^2 + 2D + 1) = 3e^{2x} - \cos x$,
(γ) $(D^2 + 4)y = x^2 + \cos x$,
(δ) $(D^3 + D^2 - D - 1)y = \cos 2x$,
(ϵ) $(D^2 + 1)^2 y = \cos x$,
(ζ) $(D^3 - D^2 + D - 1)y = \cos x$,
(η) $(D^2 - 5D + 6)y = \cos x - e^{2x}$,
(θ) $(D^3 - 2D^2 - 3D)y = 3x^2 + \sin x$,
(ι) $(D^2 - 1)^2 y = \sin x$,
(κ) $(D^2 + 3D + 2)y = e^{2x}\sin x$,
(λ) $(D^4 - 1)y = e^x \cos x$,
(μ) $(D^3 - 3D^2 + 4D - 2)y = e^x + \cos x$,
(ν) $(D^2 - 2D + 4)y = e^x \sin x$,
(o) $(D^2 + 4)y = \sin 3x + e^x + x^2$,
(π) $(D^6 + 1)y = \sin \frac{3}{2}x \sin \frac{1}{2}x$,
(ρ) $(D^3 + 1)y = e^{2x}\sin x + e^{\frac{x}{2}}\sin\frac{x\sqrt{3}}{2}$,
(σ) $(D^2 + 4)y = \sin^2 x$,
(τ) $(D^4 + 32D + 48)y = xe^{-2x} + e^{2x}\cos 2\frac{3}{2}x$.

6. If X has the form $e^{ax}X_1$, show that $I = \frac{1}{P(D)}e^{ax}X_1 = e^{ax}\frac{1}{P(D+\alpha)}X_1$. This enables the solution of equations where X_1 is a polynomial to be obtained by a short method; it also gives a way of treating equations where X is $e^{ax}\cos\beta x$ or $e^{ax}\sin\beta x$, but is not an improvement on (24); finally, combined with the second suggestion of (24), it covers the case where X is the product of a sine or cosine by a polynomial. Solve by this method, or partly by this method, (ζ) of Ex. 1; (κ), (λ), (ν), (ρ), (τ) of Ex. 5; and also

(α) $(D^2 - 2D + 1)y = x^2e^{3x}$,
(β) $(D^3 + 3D^2 + 3D + 1)y = (2 - x^2)e^{-x}$,
(γ) $(D^2 + n^2)y = x^4e^x$,
(δ) $(D^4 - 2D^3 - 3D^2 + 4D + 4)y = x^2e^x$,
(ϵ) $(D^3 - 7D - 6)y = e^{2x}(1 + x)$,
(ζ) $(D - 1)^2 y = e^x + \cos x + x^2e^x$,
(η) $(D - 1)^3 y = x - x^3e^x$,
(θ) $(D^2 + 2)y = x^2e^{3x} + e^x\cos 2x$,
(ι) $(D^3 - 1)y = xe^x + \cos^2 x$,
(κ) $(D^2 - 1)y = x\sin x + (1 + x^2)e^x$,
(λ) $(D^2 + 4)y = x\sin x$,
(μ) $(D^4 + 2D^2 + 1)y = x^2\cos ax$,
(ν) $(D^2 + 4)y = (x\sin x)^2$,
(o) $(D^2 - 2D + 4)^2 y = xe^x\cos\sqrt{3}x$.

7. Show that the substitution $x = e^t$, Ex. 9, p. 152, changes equations of the type

$$x^nD^ny + a_1x^{n-1}D^{n-1}y + \cdots + a_{n-1}xDy + a_ny = X(x) \qquad (26)$$

into equations with constant coefficients; also that $ax + b = e^t$ would make a similar simplification for equations whose coefficients were powers of $ax + b$. Hence integrate:

(α) $(x^2D^2 - xD + 2)y = x\log x$,
(β) $(x^3D^3 - x^2D^2 + 2xD - 2)y = x^3 + 3x$,
(γ) $[(2x-1)^3D^3 + (2x-1)D - 2]y = 0$,
(δ) $(x^2D^2 + 3xD + 1)y = (1 - x)^{-2}$,
(ϵ) $(x^3D^3 + xD - 1)y = x\log x$,
(ζ) $[(x + 1)^2D^2 - 4(x + 1)D + 6]y = x$,
(η) $(x^2D + 4xD + 2)y = e^x$,
(θ) $(x^3D^2 - 3x^2D + x)y = \log x \sin\log x + 1$,
(ι) $(x^4D^4 + 6x^3D^3 + 4x^2D^2 - 2xD - 4)y = x^2 + 2\cos\log x$.

8. If L be self-induction, R resistance, C capacity, i current, q charge upon the plates of a condenser, and $f(t)$ the electromotive force, then the differential equations for the circuit are

$$(\alpha)\ \frac{d^2q}{dt^2} + \frac{R}{L}\frac{dq}{dt} + \frac{q}{LC} = \frac{1}{L}f(t), \qquad (\beta)\ \frac{d^2i}{dt^2} + \frac{R}{L}\frac{di}{dt} + \frac{i}{LC} = \frac{1}{L}f'(t).$$

Solve (α) when $f(t) = e^{-ax}\sin bx$ and (β) when $f(t) = \sin bx$. Reduce the trigonometric part of the particular solution to the form $K\sin(bx + \gamma)$. Show that if R is small and b is nearly equal to $1/\sqrt{LC}$, the amplitude K is large.

98. Simultaneous linear equations with constant coefficients. If there be given two (or in general n) linear equations with constant coefficients in two (or in general n) dependent variables and one independent variable t, the symbolic method of solution may still be used to advantage. Let the equations be

$$\begin{aligned}(a_0D^n + a_1D^{n-1} + \cdots + a_n)x + (b_0D^m + b_1D^{m-1} + \cdots + b_m)y &= R(t),\\ (c_0D^p + c_1D^{p-1} + \cdots + c_p)x + (d_0D^q + d_1D^{q-1} + \cdots + d_q)y &= S(t),\end{aligned} \quad (27)$$

when there are two variables and where D denotes differentiation by t. The equations may also be written more briefly as

$$P_1(D)x + Q_1(D)y = R \quad \text{and} \quad P_2(D)x + Q_2(D)y = S.$$

The ordinary algebraic process of solution for x and y may be employed because it depends only on such laws as are satisfied equally by the symbols D, $P_1(D)$, $Q_1(D)$, and so on.

Hence the solution for x and y is found by multiplying by the appropriate coefficients and adding the equations.

$$\begin{array}{r|r|l} Q_2(D) & -P_2(D) & P_1(D)x + Q_1(D)y = R,\\ -Q_1(D) & P_1(D) & P_2(D)x + Q_2(D)y = S.\end{array}$$

Then
$$\begin{aligned}[P_1(D)\,Q_2(D) - P_2(D)\,Q_1(D)]\,x &= Q_2(D)\,R - Q_1(D)\,S,\\ [P_1(D)\,Q(D) - P_2(D)\,Q_1(D)]\,y &= P_1(D)\,S - P_2(D)\,R.\end{aligned} \quad (27')$$

It will be noticed that the coefficients by which the equations are multiplied (written on the left) are so chosen as to make the coefficients of x and y in the solved form the same in sign as in other respects. It may also be noted that the order of P and Q in the symbolic products is immaterial. By expanding the operator $P_1(D)\,Q_2(D) - P_2(D)\,Q_1(D)$ a certain polynomial in D is obtained and by applying the operators to R and S as indicated certain functions of t are obtained. Each equation, whether in x or in y, is quite of the form that has been treated in §§ 95–97.

As an example consider the solution for x and y in the case of

$$2\frac{d^2x}{dt^2} - \frac{dy}{dt} - 4x = 2t, \qquad 2\frac{dx}{dt} + 4\frac{dy}{dt} - 3y = 0;$$

or
$$(2D^2 - 4)x - Dy = 2t, \qquad 2Dx + (4D - 3)y = 0.$$

Solve
$$\begin{array}{r|r|l} 4D - 3 & -2D & (2D^2 - 4)x - Dy = 2t\\ D & 2D^2 - 4 & 2Dx + (4D - 3)y = 0.\end{array}$$

Then
$$\begin{aligned}&[(4D - 3)(2D^2 - 4) + 2D^2]\,x = (4D - 3)\,2t,\\ &[2D^2 + (2D^2 - 4)(4D - 3)]\,y = -(2D)\,2t,\end{aligned}$$

or
$$4(2D^3 - D^2 - 4D + 3)x = 8 - 6t, \qquad 4(2D^3 - D^2 - 4D + 3)y = -4.$$

The roots of the polynomial in D are 1, 1, $-1\frac{1}{2}$; and the particular solution I_x for x is $-\frac{1}{2}t$, and I_y for y is $-\frac{1}{3}$. Hence the solutions have the form

$$x = (C_1 + C_2t)\,e^t + C_3e^{-\frac{3}{2}t} - \tfrac{1}{2}t, \qquad y = (K_1 + K_2t)\,e^t + K_3e^{-\frac{3}{2}t} - \tfrac{1}{3}.$$

The arbitrary constants which are introduced into the solutions for x and y are not independent nor are they identical. *The solutions must be substituted into one of the equations to establish the necessary relations between the constants.* It will be noticed that in general the order of the equation in D for x and for y is the sum of the orders of the highest derivatives which occur in the two equations, — in this case, $3 = 2 + 1$. The order may be diminished by cancellations which occur in the formal algebraic solutions for x and y. In fact it is conceivable that the coefficient $P_1Q_2 - P_2Q_1$ of x and y in the solved equations should vanish and the solution become illusory. This case is of so little consequence in practice that it may be dismissed with the statement that the solution is then either impossible or indeterminate; that is, either there are no functions x and y of t which satisfy the two given differential equations, or there are an infinite number in each of which other things than the constants of integration are arbitrary.

To finish the example above and determine one set of arbitrary constants in terms of the other, substitute in the second differential equation.. Then

$$2\left(C_1e^t + C_2e^t + C_2te^t - \tfrac{3}{2}C_3e^{-\frac{3}{2}t} - \tfrac{1}{2}\right) + 4\left(K_1e^t + K_2e^t + K_2te^t - \tfrac{3}{2}K_3e^{-\frac{3}{2}t}\right)$$
$$- 3\left(K_1e^t + K_2te^t + K_3e^{-\frac{3}{2}t} - \tfrac{1}{3}\right) = 0,$$

or $$e^t(2C_1 + 2C_2 + K_1 + K_2) + te^t(2C_2 + K_2) - 3e^{-\frac{3}{2}t}(C_3 + 3K_3) = 0.$$

As the terms e^t, te^t, $e^{-\frac{3}{2}t}$ are independent, the linear relation between them can hold only if each of the coefficients vanishes. Hence

$$C_3 + 3K_3 = 0, \qquad 2C_2 + K_2 = 0, \qquad 2C_1 + 2C_2 + K_1 + K_2 = 0,$$

and $$C_3 = -3K_3, \qquad 2C_2 = -K_2, \qquad 2C_1 = -K_1.$$

Hence $x = (C_1 + C_2t)e^t - 3K_3e^{-\frac{3}{2}t} - \frac{1}{2}t, \qquad y = -2(C_1 + C_2t)e^t + K_3e^{-\frac{3}{2}t} - \frac{1}{3}$

are the finished solutions, where C_1, C_2, K_3 are three arbitrary constants of integration and might equally well be denoted by C_1, C_2, C_3, or K_1, K_2, K_3.

99. One of the most important applications of the theory of simultaneous equations with constant coefficients is to *the theory of small vibrations about a state of equilibrium in a conservative* dynamical system.* If $q_1, q_2, \cdots, q_n$ are n coördinates (see Exs. 19–20, p. 112) which specify the position of the system measured relatively

* The potential energy V is defined as $-dV = dW = Q_1dq_1 + Q_2dq_2 + \cdots + Q_ndq_n$, where
$$Q_i = X_1\frac{\partial x_1}{\partial q_i} + Y_1\frac{\partial y_1}{\partial q_i} + Z_1\frac{\partial z_1}{\partial q_i} + \cdots + X_n\frac{\partial x_n}{\partial q_i} + Y_n\frac{\partial y_n}{\partial q_i} + Z_n\frac{\partial z_n}{\partial q_i}.$$
This is the immediate extension of Q_1 as given in Ex. 19, p. 112. Here dW denotes the differential of work and $dW = \Sigma\mathbf{F}_i\cdot d\mathbf{r}_i = \Sigma(X_idx_i + Y_idy_i + Z_idz_i)$. To find Q_i it is generally quickest to compute dW from this relation with dx_i, dy_i, dz_i expressed in terms of the differentials $dq_1, \cdots, dq_n$. The generalized forces Q_i are then the coefficients of dq_i. If there is to be a potential V, the differential dW must be exact. It is frequently easy to find V directly in terms of $q_1, \cdots, q_n$ rather than through the mediation of $Q_1, \cdots, Q_n$; when this is not so, it is usually better to leave the equations in the form $\dfrac{d}{dt}\dfrac{\partial T}{\partial \dot{q}_i} - \dfrac{\partial T}{\partial q_i} = Q_i$ rather than to introduce V and L.

to a position of stable equilibrium in which all the q's vanish, the development of the potential energy by Maclaurin's Formula gives

$$V(q_1, q_2, \cdots, q_n) = V_0 + V_1(q_1, q_2, \cdots, q_n) + V_2(q_1, q_2, \cdots, q_n) + \cdots,$$

where the first term is constant, the second is linear, and the third is quadratic, and where the supposition that the q's take on only small values, owing to the restriction to small vibrations, shows that each term is infinitesimal with respect to the preceding. Now the constant term may be neglected in any expression of potential energy. As the position when all the q's are 0 is assumed to be one of equilibrium, the forces

$$Q_1 = -\frac{\partial V}{\partial q_1}, \quad Q_2 = -\frac{\partial V}{\partial q_2}, \quad \cdots, \quad Q_n = -\frac{\partial V}{\partial q_n}$$

must all vanish when the q's are 0. This shows that the coefficients, $(\partial V/\partial q_i)_0 = 0$, of the linear expression are all zero. Hence the first term in the expansion is the quadratic term, and relative to it the higher terms may be disregarded. As the position of equilibrium is stable, the system will tend to return to the position where all the q's are 0 when it is slightly displaced from that position. It follows that the quadratic expression must be definitely positive.

The kinetic energy is always a quadratic function of the velocities $\dot{q}_1, \dot{q}_2, \cdots, \dot{q}_n$ with coefficients which may be functions of the q's. If each coefficient be expanded by the Maclaurin Formula and only the first or constant term be retained, the kinetic energy becomes a quadratic function with constant coefficients. Hence the Lagrangian function (cf. § 160)

$$L = T - V = T(\dot{q}_1, \dot{q}_2, \cdots, \dot{q}_n) - V(q_1, q_2, \cdots, q_n),$$

when substituted in the formulas for the motion of the system, gives

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_1} - \frac{\partial L}{\partial q_1} = 0, \quad \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_2} - \frac{\partial L}{\partial q_2} = 0, \quad \cdots, \quad \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_n} - \frac{\partial L}{\partial q_n} = 0,$$

a set of equations of the second order with constant coefficients. The equations moreover involve the operator D only through its square, and the roots of the equation in D must be either real or pure imaginary. The pure imaginary roots introduce trigonometric functions in the solution and represent vibrations. If there were real roots, which would have to occur in pairs, the positive root would represent a term of exponential form which would increase indefinitely with the time, — a result which is at variance both with the assumption of stable equilibrium and with the fact that the energy of the system is constant.

When there is friction in the system, the forces of friction are supposed to vary with the velocities for small vibrations. In this case there exists a dissipative function $F(\dot{q}_1, \dot{q}_2, \cdots, \dot{q}_n)$ which is quadratic in the velocities and may be assumed to have constant coefficients. The equations of motion of the system then become

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_1} - \frac{\partial L}{\partial q_1} + \frac{\partial F}{\partial \dot{q}_1} = 0, \quad \cdots, \quad \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_n} - \frac{\partial L}{\partial q_n} + \frac{\partial F}{\partial \dot{q}_n} = 0,$$

which are still linear with constant coefficients but involve first powers of the operator D. It is physically obvious that the roots of the equation in D must be negative if real, and must have their real parts negative if the roots are complex; for otherwise the energy of the motion would increase indefinitely with the time, whereas it is known to be steadily dissipating its initial energy. It may be added that if, in addition to the internal forces arising from the potential V and the frictional forces arising from the dissipative function F, there are other forces impressed on the system, these forces would remain to be inserted upon the right-hand side of the equations of motion just given.

The fact that the equations for small vibrations lead to equations with constant coefficients by neglecting the higher powers of the variables gives the important physical theorem of the superposition of small vibrations. The theorem is: If with a certain set of initial conditions, a system executes a certain motion; and if with a different set of initial conditions taken at the same initial time, the system executes a second motion; then the system may execute the motion which consists of merely adding or superposing these motions at each instant of time; and in particular this combined motion will be that which the system would execute under initial conditions which are found by simply adding the corresponding values in the two sets of initial conditions. This theorem is of course a mere corollary of the linearity of the equations.

EXERCISES

1. Integrate the following systems of equations:

(α) $Dx - Dy + x = \cos t$, $\quad D^2x - Dy + 3x - y = e^{2t}$,
(β) $3Dx + 3x + 2y = e^t$, $\quad 4x - 3Dy + 3y = 3t$,
(γ) $D^2x - 3x - 4y = 0$, $\quad D^2y + x + y = 0$,
(δ) $\dfrac{dx}{y - 7x} = \dfrac{-dy}{2x + 5y} = dt$, $\quad$ (ϵ) $-dt = \dfrac{dx}{3x + 4y} = \dfrac{dy}{2x + 5y}$,
(ζ) $tDx + 2(x - y) = 1$, $\quad tDy + x + 5y = t$,
(η) $Dx = ny - mz$, $\quad Dy = lz - nx$, $\quad Dz = mx - ly$,
(θ) $D^2x - 3x - 4y + 3 = 0$, $\quad D^2y + x - 8y + 5 = 0$,
(ι) $D^4x - 4D^3y + 4D^2x - x = 0$, $\quad D^4y - 4D^3x + 4D^2y - y = 0$.

2. A particle vibrates without friction upon the inner surface of an ellipsoid. Discuss the motion. Take the ellipsoid as

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{(z-c)^2}{c^2} = 1; \quad \text{then} \quad x = C\sin\left(\frac{\sqrt{cg}}{a}t + C_1\right), \quad y = K\sin\left(\frac{\sqrt{cg}}{b}t + K_1\right).$$

3. Same as Ex. 2 when friction varies with the velocity.

4. Two heavy particles of equal mass are attached to a light string, one at the middle, one at one end, and are suspended by attaching the other end of the string to a fixed point. If the particles are slightly displaced and the oscillations take place without friction in a vertical plane containing the fixed point, discuss the motion.

5. If there be given two electric circuits without capacity, the equations are

$$L_1\frac{di_1}{dt} + M\frac{di_2}{dt} + R_1i_1 = E_1, \qquad L_2\frac{di_2}{dt} + M\frac{di_1}{dt} + R_2i_2 = E_2,$$

where i_1, i_2 are the currents in the circuits, L_1, L_2 are the coefficients of self-induction, R_1, R_2 are the resistances, and M is the coefficient of mutual induction. (α) Integrate the equations when the impressed electromotive forces E_1, E_2 are zero in both circuits. (β) Also when $E_2 = 0$ but $E_1 = \sin pt$ is a periodic force. (γ) Discuss the cases of loose coupling, that is, where M^2/L_1L_2 is small; and the case of close coupling, that is, where M^2/L_1L_2 is nearly unity. What values for p are especially noteworthy when the damping is small?

6. If the two circuits of Ex. 5 have capacities C_1, C_2 and if q_1, q_2 are the charges on the condensers so that $i_1 = dq_1/dt$, $i_2 = dq_2/dt$ are the currents, the equations are

$$L_1\frac{d^2q_1}{dt^2} + M\frac{d^2q_2}{dt^2} + R_1\frac{dq_1}{dt} + \frac{q_1}{C_1} = E_1, \qquad L_2\frac{d^2q_2}{dt^2} + M\frac{d^2q_1}{dt^2} + R_2\frac{dq_2}{dt} + \frac{q_2}{C_2} = E_2.$$

Integrate when the resistances are negligible and $E_1 = E_2 = 0$. If $T_1 = 2\pi\sqrt{C_1L_1}$ and $T_2 = 2\pi\sqrt{C_2L_2}$ are the periods of the individual separate circuits and $\Theta = 2\pi M\sqrt{C_1C_2}$, and if $T_1 = T_2$, show that $\sqrt{T^2+\Theta^2}$ and $\sqrt{T^2-\Theta^2}$ are the independent periods in the coupled circuits.

7. A uniform beam of weight 6 lb. and length 2 ft. is placed orthogonally across a rough horizontal cylinder 1 ft. in diameter. To each end of the beam is suspended a weight of 1 lb. upon a string 1 ft. long. Solve the motion produced by giving one of the weights a slight horizontal velocity. Note that in finding the kinetic energy of the beam, the beam may be considered as rotating about its middle point (§ 39).

CHAPTER IX

ADDITIONAL TYPES OF ORDINARY EQUATIONS

100. Equations of the first order and higher degree. The *degree* of a differential equation is defined as the degree of the derivative of highest order which enters in the equation. In the case of the equation $\Psi(x, y, y') = 0$ of the first order, the degree will be the degree of the equation in y'. From the idea of the lineal element (§ 85) it appears that if the degree of Ψ in y' is n, there will be n lineal elements through each point (x, y). Hence it is seen that there are n curves, which are compounded of these elements, passing through each point. It may be pointed out that equations such as $y' = x\sqrt{1 + y^2}$, which are apparently of the first degree in y', are really of higher degree if the multiple value of the functions, such as $\sqrt{1 + y^2}$, which enter in the equation, is taken into consideration; the equation above is replaceable by $y'^2 = x^2 + x^2y^2$, which is of the second degree and without any multiple valued function.*

First suppose that *the differential equation*

$$\Psi(x, y, y') = [y' - \psi_1(x, y)] \times [y' - \psi_2(x, y)] \cdots = 0 \qquad (1)$$

may be solved for y'. It then becomes equivalent to the set

$$y' - \psi_1(x, y) = 0, \qquad y' - \psi_2(x, y) = 0, \cdots \qquad (1')$$

of equations each of the first order, and each of these may be treated by the methods of Chap. VIII. Thus a set of integrals †

$$F_1(x, y, C) = 0, \qquad F_2(x, y, C) = 0, \cdots \qquad (2)$$

may be obtained, and the product of these separate integrals

$$F(x, y, C) = F_1(x, y, C) \cdot F_2(x, y, C) \cdots = 0 \qquad (2')$$

is the complete solution of the original equation. Geometrically speaking, each integral $F_i(x, y, C) = 0$ represents a family of curves and the product represents all the families simultaneously.

* It is therefore apparent that the idea of degree as applied in practice is somewhat indefinite.

† The same constant C or any desired function of C may be used in the different solutions because C is an arbitrary constant and no specialization is introduced by its repeated use in this way.

As an example consider $y'^2 + 2\,y'y\cot x = y^2$. Solve.

$$y'^2 + 2\,y'y\cot x + y^2\cot^2 x = y^2(1+\cot^2 x) = y^2\csc^2 x,$$

and
$$(y' + y\cot x - y\csc x)(y' + y\cot x + y\csc x) = 0.$$

These equations both come under the type of variables separable. Integrate.

$$\frac{dy}{y} = \frac{1-\cos x}{\sin x}dx = -\frac{d\cos x}{1+\cos x}, \qquad y(1+\cos x) = C,$$

and
$$\frac{dy}{y} = -\frac{1+\cos x}{\sin x}dx = \frac{d\cos x}{1-\cos x}, \qquad y(1-\cos x) = C.$$

Hence
$$[y(1+\cos x) + C][y(1-\cos x) + C] = 0$$

is the solution. It may be put in a different form by multiplying out. Then

$$y^2\sin^2 x + 2\,Cy + C^2 = 0.$$

If the equation cannot be solved for y' or if the equations resulting from the solution cannot be integrated, this first method fails. In that case *it may be possible to solve for y or for x* and treat the equation by differentiation. Let $y' = p$. Then if

$$y = f(x, p), \qquad \frac{dy}{dx} = p = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial p}\frac{dp}{dx}. \tag{3}$$

The equation thus found by differentiation is a differential equation of the first order in dp/dx and it may be solved by the methods of Chap. VIII to find $F(p, x, C) = 0$. The two equations

$$y = f(x, p) \quad \text{and} \quad F(p, x, C) = 0 \tag{3'}$$

may be regarded as defining x and y parametrically in terms of p, or p may be eliminated between them to determine the solution in the form $\Omega(x, y, C) = 0$ if this is more convenient. If the given differential equation had been solved for x, then

$$x = f(y, p) \quad \text{and} \quad \frac{dx}{dy} = \frac{1}{p} = \frac{\partial f}{\partial y} + \frac{\partial f}{\partial p}\frac{dp}{dy}. \tag{4}$$

The resulting equation on the right is an equation of the first order in dp/dy and may be treated in the same way.

As an example take $xp^2 - 2\,yp + ax = 0$ and solve for y. Then

$$2\,y = xp + \frac{ax}{p}, \qquad 2\frac{dy}{dx} = 2\,p = p + x\frac{dp}{dx} - \frac{ax}{p^2}\frac{dp}{dx} + \frac{a}{p},$$

or
$$\frac{x}{p}\left[p - \frac{a}{p}\right]\frac{dp}{dx} + \left(\frac{a}{p} - p\right) = 0, \quad \text{or} \quad xdp - pdx = 0.$$

The solution of this equation is $x = Cp$. The solution of the given equation is

$$2\,y = xp + \frac{ax}{p}, \qquad x = Cp$$

when expressed parametrically in terms of p. If p be eliminated, then

$$2\,y = \frac{x^2}{C} + aC \qquad \text{parabolas.}$$

As another example take $p^2y + 2px = y$ and solve for x. Then

$$2x = y\left(\frac{1}{p} - p\right), \quad 2\frac{dx}{dy} = \frac{2}{p} = \frac{1}{p} - p + y\left(-\frac{1}{p^2} - 1\right)\frac{dp}{dy},$$

or
$$\frac{1}{p} + p + y\left(\frac{1}{p^2} + 1\right)\frac{dp}{dy} = 0, \quad \text{or} \quad ydp + pdy = 0.$$

The solution of this is $py = C$ and the solution of the given equation is

$$2x = y\left(\frac{1}{p} - p\right), \quad py = C, \quad \text{or} \quad y^2 = 2Cx + C^2.$$

Two special types of equation may be mentioned in addition, although their method of solution is a mere corollary of the methods already given in general. They are the equation *homogeneous* in (x, y) and *Clairaut's* equation. The general form of the homogeneous equation is $\Psi(p, y/x) = 0$. This equation may be solved as

$$p = \psi\left(\frac{y}{x}\right) \qquad \text{or as} \quad \frac{y}{x} = f(p), \qquad y = xf(p); \tag{5}$$

and in the first case is treated by the methods of Chap. VIII, and in the second by the methods of this article. Which method is chosen rests with the solver. The Clairaut type of equation is

$$y = px + f(p) \tag{6}$$

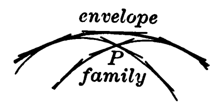
and comes directly under the methods of this article. It is especially noteworthy, however, that on differentiating with respect to x the resulting equation is

$$[x + f'(p)]\frac{dp}{dx} = 0 \quad \text{or} \quad \frac{dp}{dx} = 0. \tag{6'}$$

Hence the solution for p is $p = C$, and thus $y = Cx + f(C)$ is the solution for the Clairaut equation and represents a family of straight lines. The rule is merely to substitute C in place of p. This type occurs very frequently in geometric applications either directly or in a disguised form requiring a preliminary change of variable.

101. To this point the only solution of the differential equation $\Psi(x, y, p) = 0$ which has been considered is the *general solution* $F(x, y, C) = 0$ containing an arbitrary constant. If a special value, say 2, is given to C, the solution $F(x, y, 2) = 0$ is called a *particular solution.* It may happen that the arbitrary constant C enters into the expression $F(x, y, C) = 0$ in such a way that when C becomes positively infinite (or negatively infinite) the curve $F(x, y, C) = 0$ approaches a definite limiting position which is a solution of the differential equation; such solutions are called *infinite solutions.* In addition to these types of solution which naturally group themselves in connection with the general solution, there is often a solution of a different kind which is known as the *singular solution.* There are several different definitions for the singular solution. That which will be adopted here is: *A singular solution is the envelope of the family of curves defined by the general solution.*

The consideration of the lineal elements (§ 85) will show how it is that the envelope (§ 65) of the family of particular solutions which constitute the general solution is itself a solution of the equation. For consider the figure, which represents the particular solutions broken up into their lineal elements. Note that the envelope is made up of those lineal elements, one taken from each particular solution, which are at the points of contact of the envelope with the curves of the family. It is seen that the envelope is a curve all of whose lineal elements satisfy the equation $\Psi(x, y, p) = 0$ for the reason that they lie upon solutions of the equation. Now any curve whose lineal elements satisfy the equation is by definition a solution of the equation; and so the envelope must be a solution. It might conceivably happen that the family $F(x, y, C) = 0$ was so constituted as to envelope one of its own curves. In that case that curve would be both a particular and a singular solution.



If the general solution $F(x, y, C) = 0$ of a given differential equation is known, the singular solution may be found according to the rule for finding envelopes (§ 65) by eliminating C from

$$F(x, y, C) = 0 \quad \text{and} \quad \frac{\partial}{\partial C} F(x, y, C) = 0. \tag{7}$$

It should be borne in mind that in the eliminant of these two equations there may occur some factors which do not represent envelopes and which must be discarded from the singular solution. If only the singular solution is desired and the general solution is not known, this method is inconvenient. In the case of Clairaut's equation, however, where the solution is known, it gives the result immediately as that obtained by eliminating C from the two equations

$$y = Cx + f(C) \quad \text{and} \quad 0 = x + f'(C). \tag{8}$$

It may be noted that as $p = C$, the second of the equations is merely the factor $x + f'(p) = 0$ discarded from (6′). The singular solution may therefore be found by eliminating p between the given Clairaut equation and the discarded factor $x + f'(p) = 0$.

A reëxamination of the figure will suggest a means of finding the singular solution without integrating the given equation. For it is seen that when two neighboring curves of the family intersect in a point P

near the envelope, then through this point there are two lineal elements which satisfy the differential equation. These two lineal elements have nearly the same direction, and indeed the nearer the two neighboring curves are to each other the nearer will their intersection lie to the envelope and the nearer will the two lineal elements approach coincidence with each other and with the element upon the envelope at the point of contact. Hence for all points (x, y) on the envelope the equation $\Psi(x, y, p) = 0$ of the lineal elements must have *double roots for p*. Now if an equation has double roots, the derivative of the equation must have a root. Hence the requirement that the two equations

$$\psi(x, y, p) = 0 \quad \text{and} \quad \frac{\partial}{\partial p}\psi(x, y, p) = 0 \tag{9}$$

have a common solution for p will insure that the first has a double root for p; and the points (x, y) which satisfy these equations simultaneously must surely include all the points of the envelope. The rule for finding the singular solution is therefore: *Eliminate p from the given differential equation and its derivative with respect to p*, that is, from (9). The result should be tested.

If the equation $xp^2 - 2yp + ax = 0$ treated above be tried for a singular solution, the elimination of p is required between the two equations

$$xp^2 - 2yp + ax = 0 \quad \text{and} \quad xp - y = 0.$$

The result is $y^2 = ax^2$, which gives a pair of lines through the origin. The substitution of $y = \pm\sqrt{a}x$ and $p = \pm\sqrt{a}$ in the given equation shows at once that $y^2 = ax^2$ satisfies the equation. Thus $y^2 = ax^2$ is a singular solution. The same result is found by finding the envelope of the general solution given above. It is clear that in this case the singular solution is not a particular solution, as the particular solutions are parabolas.

If the elimination had been carried on by Sylvester's method, then

$$\begin{vmatrix} 0 & x & -y \\ x & -2y & a \\ x & -y & 0 \end{vmatrix} = -x(y^2 - ax^2) = 0;$$

and the eliminant is the product of two factors $x = 0$ and $y^2 - ax^2 = 0$, of which the second is that just found and the first is the y-axis. As the slope of the y-axis is infinite, the substitution in the equation is hardly legitimate, and the equation can hardly be said to be satisfied. The occurrence of these extraneous factors in the eliminant is the real reason for the necessity of testing the result to see if it actually represents a singular solution. These extraneous factors may represent a great variety of conditions. Thus in the case of the equation $p^2 + 2yp\cot x = y^2$ previously treated, the elimination gives $y^2 \csc^2 x = 0$, and as $\csc x$ cannot vanish, the result reduces to $y^2 = 0$, or the x-axis. As the slope along the x-axis is 0 and y is 0, the equation is clearly satisfied. Yet the line $y = 0$ is *not* the envelope of the general solution; for the curves of the family touch the line only at the points $n\pi$. It is a particular solution and corresponds to $C = 0$. There is no singular solution.

Many authors use a great deal of time and space discussing just what may and what may not occur among the extraneous loci and how many times it may occur. The result is a considerable number of statements which in their details are either grossly incomplete or glaringly false or both (cf. §§ 65–67). The rules here given for finding singular solutions should not be regarded in any other light than as leading to some expressions which are to be examined, the best way one can, to find out whether or not they are singular solutions. One curve which may appear in the elimination of p and which deserves a note is the *tac-locus* or locus of points of tangency of the particular solutions with each other. Thus in the system of circles $(x-C)^2+y^2=r^2$ there may be found two which are tangent to each other at any assigned point of the x-axis. This tangency represents two coincident lineal elements and hence may be expected to occur in the elimination of p between the differential equation of the family and its derivative with respect to p; but not in the eliminant from (7).

EXERCISES

1. Integrate the following equations by solving for $p = y'$:

(α) $p^2-6p+5=0$, (β) $p^3-(2x+y^2)p^2+(x^2-y^2+2xy^2)p-(x^2-y^2)y^2=0$,
(γ) $xp^2-2yp-x=0$, (δ) $p^3(x+2y)+3p^2(x+y)+p(y+2x)=0$,
(ϵ) $y^2+p^2=1$, (ζ) $p^2-ax^3=0$, (η) $p=(a-x)\sqrt{1+p^2}$.

2. Integrate the following equations by solving for y or x:

(α) $4xp^2+2xp-y=0$, (β) $y=-xp+x^4p^2$, (γ) $p+2xy-x^2-y^2=0$,
(δ) $2px-y+\log p=0$, (ϵ) $x-yp=ap^2$, (ζ) $y=x+a\tan^{-1}p$,
(η) $x=y+a\log p$, (θ) $x+py(2p^2+3)=0$, (ι) $a^2yp^2-2xp+y=0$,
(κ) $p^3-4xyp+8y^2=0$, (λ) $x=p+\log p$, (μ) $p^2(x^2+2ax)=a^2$.

3. Integrate these equations [substitutions suggested in (ι) and (κ)]:

(α) $xy^2(p^2+2)=2py^3+x^3$, (β) $(nx+py)^2=(1+p^2)(y^2+nx^2)$,
(γ) $y^2+xyp-x^2p^2=0$, (δ) $y=yp^2+2px$,
(ϵ) $y=px+\sin^{-1}p$, (ζ) $y=p(x-b)+a/p$,
(η) $y=px+p(1-p^2)$, (θ) $y^2-2pxy-1=p^2(1-x^2)$,
(ι) $4e^{2y}p^2+2xp-1=0$, $z=e^{2y}$, (κ) $y=2px+y^2p^3$, $y^2=z$,
(λ) $4e^{2y}p^2+2e^{2x}p-e^x=0$, ($\mu$) $x^2(y-px)=yp^2$.

4. Treat these equations by the p method (9) to find the singular solutions. Also solve and treat by the C method (7). Sketch the family of solutions and examine the significance of the extraneous factors as well as that of the factor which gives the singular solution:

(α) $p^2y+p(x-y)-x=0$, (β) $p^2y^2\cos^2\alpha-2pxy\sin^2\alpha+y^2-x^2\sin^2\alpha=0$,
(γ) $4xp^2=(3x-a)^2$, (δ) $yp^2x(x-a)(x-b)=[3x^2-2x(a+b)+ab]^2$,
(ϵ) $p^2+xp-y=0$, (ζ) $8a(1+p)^3=27(x+y)(1-p)^3$,
(η) $x^3p^2+x^2yp+a^3=0$, (θ) $y(3-4y)^2p^2=4(1-y)$.

5. Examine sundry of the equations of Exs. 1, 2, 3, for singular solutions.

6. Show that the solution of $y=x\phi(p)+f(p)$ is given parametrically by the given equation and the solution of the linear equation:

$$\frac{dx}{dp}+x\frac{\phi'(p)}{\phi(p)-p}=\frac{f'(p)}{p-\phi(p)}.\qquad \text{Solve}\quad (\alpha)\ y=mxp+n(1+p^3)^{\frac{1}{3}},$$

(β) $y=x(p+a\sqrt{1+p^2})$, (γ) $x=yp+ap^2$, (δ) $y=(1+p)x+p^2$.

7. As any straight line is $y = mx + b$, any family of lines may be represented as $y = mx + f(m)$ or by the Clairaut equation $y = px + f(p)$. Show that the orthogonal trajectories of any family of lines leads to an equation of the type of Ex. 6. The same is true of the trajectories at any constant angle. Express the equations of the following systems of lines in the Clairaut form, write the equations of the orthogonal trajectories, and integrate:

(α) tangents to $x^2 + y^2 = 1$, (β) tangents to $y^2 = 2ax$,
(γ) tangents to $y^2 = x^3$, (δ) normals to $y^2 = 2ax$,
(ϵ) normals to $y^2 = x^3$, (ζ) normals to $b^2x^2 + a^2y^2 = a^2b^2$.

8. The *evolute* of a given curve is the locus of the center of curvature of the curve, or, what amounts to the same thing, it is the envelope of the normals of the given curve. If the Clairaut equation of the normals is known, the evolute may be obtained as its singular solution. Thus find the evolutes of

(α) $y^2 = 4ax$, (β) $2xy = a^2$, (γ) $x^{\frac{2}{3}} + y^{\frac{2}{3}} = a^{\frac{2}{3}}$,
(δ) $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 1$, ($\epsilon$) $y^2 = \dfrac{x^3}{2a - x}$, ($\zeta$) $y = \frac{1}{2}(e^x + e^{-x})$.

9. The *involutes* of a given curve are the curves which cut the tangents of the given curve orthogonally, or, what amounts to the same thing, they are the curves which have the given curve as the locus of their centers of curvature. Find the involutes of

(α) $x^2 + y^2 = a^2$, (β) $y^2 = 2mx$, (γ) $y = a\cosh(x/a)$.

10. As any curve is the envelope of its tangents, it follows that when the curve is described by a property of its tangents the curve may be regarded as the singular solution of the Clairaut equation of its tangent lines. Determine thus what curves have these properties:

(α) length of the tangent intercepted between the axes is l,
(β) sum of the intercepts of the tangent on the axes is c,
(γ) area between the tangent and axes is the constant k^2,
(δ) product of perpendiculars from two fixed points to tangent is k^2,
(ϵ) product of ordinates from two points of x-axis to tangent is k^2.

11. From the relation $\dfrac{dF}{dn} = \mu\sqrt{M^2 + N^2}$ of Proposition 3, p. 212, show that as the curve $F = C$ is moving tangentially to itself along its envelope, the singular solution of $Mdx + Ndy = 0$ may be expected to be found in the equation $1/\mu = 0$; also the infinite solutions. Discuss the equation $1/\mu = 0$ in the following cases:

(α) $\sqrt{1 - y^2}\,dx = \sqrt{1 - x^2}\,dy$, ($\beta$) $xdx + ydy = \sqrt{x^2 + y^2 - a^2}\,dy$.

102. Equations of higher order. In the treatment of special problems (§ 82) it was seen that the substitutions

$$\frac{dy}{dx} = p, \quad \frac{d^2y}{dx^2} = \frac{dp}{dx} \quad \text{or} \quad \frac{d^2y}{dx^2} = p\frac{dp}{dy} \tag{10}$$

rendered the differential equations integrable by reducing them to integrable equations of the first order. These substitutions or others like them are useful in treating certain cases of the differential equation

$\Psi(x, y, y', y'', \cdots, y^{(n)}) = 0$ of the nth order, namely, when one of the variables and perhaps some of the derivatives of lowest order do not occur in the equation.

In case $$\Psi\left(x, \frac{d^i y}{dx^i}, \frac{d^{i+1}y}{dx^{i+1}}, \cdots, \frac{d^n y}{dx^n}\right) = 0, \tag{11}$$

y and the first $i - 1$ derivatives being absent, substitute

$$\frac{d^i y}{dx^i} = q \quad \text{so that} \quad \Psi\left(x, q, \frac{dq}{dx}, \cdots, \frac{d^{n-i}q}{dx^{n-i}}\right) = 0. \tag{11'}$$

The original equation is therefore replaced by one of lower order. If the integral of this be $F(x, q) = 0$, which will of course contain $n - i$ arbitrary constants, the solution for q gives

$$q = f(x) \quad \text{and} \quad y = \int \cdots \int f(x)\,(dx)^i. \tag{12}$$

The solution has therefore been accomplished. If it were more convenient to solve $F(x, q) = 0$ for $x = \phi(q)$, the integration would be

$$y = \int \cdots \int q\,(dx)^i = \int \cdots \int q\,[\phi'(q)\,dq]^i; \tag{12'}$$

and this equation with $x = \phi(q)$ would give a parametric expression for the integral of the differential equation.

In case $$\Psi\left(y, \frac{dy}{dx}, \frac{d^2y}{dx^2}, \cdots, \frac{d^n y}{dx^n}\right) = 0, \tag{13}$$

x being absent, substitute p and regard p as a function of y. Then

$$\frac{dy}{dx} = p, \qquad \frac{d^2y}{dx^2} = p\frac{dp}{dy}, \qquad \frac{d^3y}{dx^3} = p\frac{d}{dy}\left(p\frac{dp}{dy}\right), \cdots$$

and $$\Psi_1\left(y, p, \frac{dp}{dy}, \cdots, \frac{d^{n-1}p}{dy^{n-1}}\right) = 0. \tag{13'}$$

In this way the order of the equation is lowered by unity. If this equation can be integrated as $F(y, p) = 0$, the last step in the solution may be obtained either directly or parametrically as

$$p = f(y), \qquad \int \frac{dy}{f(y)} = x \tag{14}$$

or $$y = \phi(p), \qquad x = \int \frac{dy}{p} = \int \frac{\phi'(p)\,dp}{p}. \tag{14'}$$

It is no particular simplification in this case to have some of the lower derivatives of y absent from $\Psi = 0$, because in general the lower derivatives of p will none the less be introduced by the substitution that is made.

As an example consider $\left(x\dfrac{d^3y}{dx^3}-\dfrac{d^2y}{dx^2}\right)^2=\left(\dfrac{d^3y}{dx^3}\right)^2+1,$

which is $$\left(x\frac{dq}{dx}-q\right)^2=\left(\frac{dq}{dx}\right)^2+1 \quad \text{if} \quad q=\frac{d^2y}{dx^2}.$$

Then $$q=x\frac{dq}{dx}\pm\sqrt{\left(\frac{dq}{dx}\right)^2+1} \quad \text{and} \quad q=C_1x\pm\sqrt{C_1^2+1};$$

for the equation is a Clairaut type. Hence, finally,

$$y=\iint\left[C_1x\pm\sqrt{C_1^2+1}\right](dx)^2=\tfrac{1}{6}C_1x^3\pm\tfrac{1}{2}x^2\sqrt{C_1^2+1}+C_2x+C_3.$$

As another example consider $y''-y'^2=y^2\log y$. This becomes

$$p\frac{dp}{dy}-p^2=y^2\log y \quad \text{or} \quad \frac{d(p^2)}{dy}-2p^2=2y^2\log y.$$

The equation is linear in p^2 and has the integrating factor e^{-2y}.

$$\frac{1}{2}p^2e^{-2y}=\int y^2e^{-2y}\log y\,dy, \qquad \frac{1}{\sqrt{2}}p=\left[e^{2y}\int y^2e^{-2y}\log y\,dy\right]^{\frac{1}{2}},$$

and $$\int\frac{dy}{\left[e^{2y}\int y^2e^{-2y}\log y\,dy\right]^{\frac{1}{2}}}=\sqrt{2}\,x.$$

The integration is therefore reduced to quadratures and becomes a problem in ordinary integration.

If an equation is *homogeneous with respect to y and its derivatives*, that is, if the equation is multiplied by a power of k when y is replaced by ky, the order of the equation may be lowered by the substitution $y=e^z$ and by taking z' as the new variable. If the equation is *homogeneous with respect to x and dx*, that is, if the equation is multiplied by a power of k when x is replaced by kx, the order of the equation may be reduced by the substitution $x=e^t$. The work may be simplified (Ex. 9, p. 152) by the use of

$$D_x^ny=e^{-nt}D_t(D_t-1)\cdots(D_t-n+1)y. \tag{15}$$

If the equation is *homogeneous with respect to x and y and the differentials dx, dy, d^2y, ⋯*, the order may be lowered by the substitution $x=e^t$, $y=e^tz$, where it may be recalled that

$$\begin{aligned}D_x^ny&=e^{-nt}D_t(D_t-1)\cdots(D_t-n+1)y\\&=e^{-(n-1)t}(D_t+1)D_t\cdots(D_t-n+2)z.\end{aligned} \tag{15'}$$

Finally, if the equation is *homogeneous with respect to x considered of dimensions* 1, *and y considered of dimensions m*, that is, if the equation is multiplied by a power of k when kx replaces x and k^my replaces y, the substitution $x=e^t$, $y=e^{mt}z$ will lower the degree of the equation. It may be recalled that

$$D_x^ny=e^{(m-n)t}(D_t+m)(D_t+m-1)\cdots(D_t+m-n+1)z. \tag{15''}$$

Consider $xyy'' - xy'^2 = yy' + bxy'^2/\sqrt{a^2 - x^2}$. If in this equation y be replaced by ky so that y' and y'' are also replaced by ky' and ky'', it appears that the equation is merely multiplied by k^2 and is therefore homogeneous of the first sort mentioned. Substitute

$$y = e^z, \qquad y' = e^z z', \qquad y'' = e^z(z'' + z'^2).$$

Then e^{2z} will cancel from the whole equation, leaving merely

$$xz'' = z' + bxz'^2/\sqrt{a^2 - x^2} \quad \text{or} \quad \frac{xdz'}{z'^2} - \frac{1}{z'}dx = \frac{bxdx}{\sqrt{a^2 - x^2}}.$$

The equation in the first form is Bernoulli; in the second form, exact. Then

$$\frac{x}{z'} = b\sqrt{a^2 - x^2} + C \quad \text{and} \quad dz = \frac{xdx}{b\sqrt{a^2 - x^2} + C}.$$

The variables are separated for the last integration which will determine $z = \log y$ as a function of x.

Again consider $x^4 \dfrac{d^2y}{dx^2} = (x^3 + 2xy)\dfrac{dy}{dx} - 4y^2$. If x be replaced by kx and y by k^2y so that y' is replaced by ky' and y'' remains unchanged, the equation is multiplied by k^4 and hence comes under the fourth type mentioned above. Substitute

$$x = e^t, \qquad y = e^{2t}z, \qquad D_xy = e^t(D_t + 2)z, \qquad D_x^2y = (D_t + 2)(D_t + 1)z.$$

Then e^{4t} will cancel and leave $z'' + 2(1 - z)z' = 0$, if accents denote differentiation with respect to t. This equation lacks the independent variable t and is reduced by the substitution $z'' = z'dz'/dz$. Then

$$\frac{dz'}{dz} + 2(1 - z) = 0, \qquad z' = \frac{dz}{dt} = (1 - z)^2 + C, \qquad \frac{dz}{(1 - z^2) + C} = dt.$$

There remains only to perform the quadrature and replace z and t by x and y.

103. If the equation may be obtained by differentiation, as

$$\Psi\left(x, y, \frac{dy}{dx}, \cdots, \frac{d^ny}{dx^n}\right) = \frac{d\Omega}{dx} = \frac{\partial\Omega}{\partial x} + \frac{\partial\Omega}{\partial y}y' + \cdots + \frac{\partial\Omega}{\partial y^{(n-1)}}y^{(n)}, \quad (16)$$

it is called an *exact equation*, and $\Omega(x, y, y', \cdots, y^{(n-1)}) = C$ is an integral of $\Psi = 0$. Thus in case the equation is exact, the order may be lowered by unity. It may be noted that unless the degree of the nth derivative is 1 the equation cannot be exact. Consider

$$\Psi(x, y, y', \cdots, y^{(n)}) = \phi_1 y^{(n)} + \phi_2,$$

where the coefficient of $y^{(n)}$ is collected into ϕ_1. Now integrate ϕ_1, partially regarding only $y^{(n-1)}$ as variable so that

$$\int \phi_1 dy^{(n-1)} = \Omega_1, \qquad \frac{d}{dx}\Omega_1 = \frac{\partial\Omega_1}{\partial x} + \cdots + \frac{\partial\Omega_1}{\partial y^{(n-2)}}y^{(n-1)} + \phi_1 y^{(n)}.$$

Then
$$\Psi - \frac{d\Omega_1}{dx} = \phi_3\left[\frac{d^{n-k}y}{dx^{n-k}}\right]^m + \phi_4.$$

That is, the expression $\Psi - \Omega_1'$ does not contain $y^{(n)}$ and may contain no derivative of order higher than $n - k$, and may be collected as indicated. Now if Ψ was an exact derivative, so must $\Psi - \Omega_1'$ be. Hence if $m \neq 1$, the conclusion is that Ψ was not exact. If $m = 1$, the process of integration may be continued to obtain Ω_2 by integrating partially with respect to $y^{(n-k-1)}$. And so on until it is shown that Ψ is not exact or until Ψ is seen to be the derivative of an expression $\Omega_1 + \Omega_2 + \cdots = C$.

As an example consider $\Psi = x^2y''' + xy'' + (2xy - 1)y' + y^2 = 0$. Then

$$\Omega_1 = \int x^2 dy'' = x^2y'', \qquad \Psi - \Omega_1' = -xy'' + (2xy-1)y' + y^2,$$

$$\Omega_2 = \int -x dy' = -xy', \qquad \Psi - \Omega_1' - \Omega_2' = 2xyy' + y^2 = (xy^2)'.$$

As the expression of the first order is an exact derivative, the result is

$$\Psi - \Omega_1' - \Omega_2' - (xy^2)' = 0; \quad \text{and} \quad \Psi_1 = x^2y'' - xy' + xy^2 - C_1 = 0$$

is the new equation. The method may be tried again.

$$\Omega_1 = \int x^2 dy' = x^2 y', \qquad \Psi_1 - \Omega_1' = -3xy' + xy^2 - C_1.$$

This is not an exact derivative and the equation $\Psi_1 = 0$ is not exact. Moreover the equation $\Psi_1 = 0$ contains both x and y and is not homogeneous of any type except when $C_1 = 0$. It therefore appears as though the further integration of the equation $\Psi = 0$ were impossible.

The method is applied with especial ease to the case of

$$X_0 \frac{d^n y}{dx^n} + X_1 \frac{d^{n-1}y}{dx^{n-1}} + \cdots + X_{n-1}\frac{dy}{dx} + X_n y - R(x) = 0, \tag{17}$$

where the coefficients are functions of x alone. This is known as the *linear equation*, the integration of which has been treated only when the order is 1 or when the coefficients are constants. The application of successive integration by parts gives

$$\Omega_1 = X_0 y^{(n-1)}, \quad \Omega_2 = (X_1 - X_0')y^{(n-2)}, \quad \Omega_3 = (X_2 - X_1' + X_0'')y^{(n-3)}, \cdots;$$

and after n such integrations there is left merely

$$(X_n - X_{n-1}' + \cdots + (-1)^{n-1}X_1 + (-1)^n X_0)y - R,$$

which is a derivative only when it is a function of x. Hence

$$X_n - X_{n-1}' + \cdots + (-1)^{n-1}X_1 + (-1)^n X_0 \equiv 0 \tag{18}$$

is the condition that the linear equation shall be exact, and

$$X_0 y^{(n-1)} + (X_1 - X_0')y^{(n-2)} + (X_2 - X_1' + X_0'')y^{(n-3)} + \cdots = \int R dx \tag{19}$$

is the first solution in case it is exact.

As an example take $y''' + y''\cos x - 2y'\sin x - y\cos x = \sin 2x$. The test

$$X_3 - X_2' + X_1'' - X_0''' = -\cos x + 2\cos x - \cos x = 0$$

is satisfied. The integral is therefore $y'' + y'\cos x - y\sin x = -\frac{1}{2}\cos 2x + C_1$. This equation still satisfies the test for exactness. Hence it may be integrated again with the result $y' + y\cos x = -\frac{1}{4}\sin 2x + C_1x + C_2$. This belongs to the linear type. The final result is therefore

$$y = e^{-\sin x}\int e^{\sin x}(C_1x + C_2)\,dx + C_3e^{-\sin x} + \tfrac{1}{2}(1 - \sin x).$$

EXERCISES

1. Integrate these equations or at least reduce them to quadratures:

(α) $2xy'''y'' = y''^2 - a^2$, (β) $(1 + x^2)y'' + 1 + y'^2 = 0$,
(γ) $y^{\text{iv}} + a^2y'' = 0$, ($\delta$) $y^{\text{v}} - m^2y''' = e^{ax}$, ($\epsilon$) $x^2y^{\text{iv}} + a^2y'' = 0$,
(ζ) $a^2y''y' = x$, (η) $xy'' + y' = 0$, (θ) $y'''y'' = 4$,
(ι) $(1 - x^2)y'' - xy' = 2$, (κ) $y^{\text{iv}} = \sqrt{y'''}$, ($\lambda$) $y'' = f(y)$,
(μ) $2(2a - y)y'' = 1 + y'^2$, (ν) $yy'' - y'^2 - y^2y' = 0$,
(o) $yy'' + y'^2 + 1 = 0$, (π) $2y'' = e^y$, (ρ) $y^3y'' = a$.

2. Carry the integration as far as possible in these cases:

(α) $x^2y'' = (mx^2y'^2 + ny^2)^{\frac{1}{2}}$, ($\beta$) $mx^3y'' = (y - xy')^2$,
(γ) $x^4y'' = (y - xy')^3$, (δ) $x^4y'' - x^3y' - x^2y'^2 + 4y^2 = 0$,
(ϵ) $x^{-2}y'' + x^{-4}y = \frac{1}{4}y'^2$, ($\zeta$) $ayy'' + by'^2 = yy'(c^2 + x^2)^{-\frac{1}{2}}$.

3. Carry the integration as far as possible in these cases:

(α) $(y^2 + x)y''' + 6yy'y'' + y'' + 2y'^2 = 0$, ($\beta$) $y'y'' - yx^2y' = xy^2$,
(γ) $x^3yy''' + 3x^3y'y'' + 9x^2yy'' + 9x^2y'^2 + 18xyy' + 3y^2 = 0$,
(δ) $y + 3xy' + 2yy'^3 + (x^2 + 2y^2y')y'' = 0$,
(ϵ) $(2x^3y' + x^2y)y'' + 4x^2y'^2 + 2xyy' = 0$.

4. Treat these linear equations:

(α) $xy'' + 2y = 2x$, (β) $(x^2 - 1)y'' + 4xy' + 2y = 2x$,
(γ) $y'' - y'\cot x + y\csc^2 x = \cos x$, ($\delta$) $(x^2 - x)y'' + (3x - 2)y' + y = 0$,
(ϵ) $(x - x^3)y''' + (1 - 5x^2)y'' - 2xy' + 2y = 6x$,
(ζ) $(x^3 + x^2 - 3x + 1)y''' + (9x^2 + 6x - 9)y'' + (18x + 6)y' + 6y = x^3$,
(η) $(x + 2)^2y''' + (x + 2)y'' + y' = 1$, ($\theta$) $x^2y'' + 3xy' + y = x$,
(ι) $(x^3 - x)y''' + (8x^2 - 3)y'' + 14xy' + 4y = 0$.

5. Note that Ex. 4 (θ) comes under the third homogeneous type, and that Ex. 4 (η) may be brought under that type by multiplying by $(x + 2)$. Test sundry of Exs. 1, 2, 3 for exactness. Show that any linear equation in which the coefficients are polynomials of degree less than the order of the derivatives of which they are the coefficients, is surely exact.

6. Sometimes, when the condition that an equation be exact is not satisfied, it is possible to find an integrating factor for the equation so that after multiplication by the factor the equation becomes exact. For linear equations try x^m. Integrate

(α) $x^5y'' + (2x^4 - x)y' - (2x^3 - 1)y = 0$, ($\beta$) $(x^2 - x^4)y'' - x^3y' - 2y = 0$.

7. Show that the equation $y'' + Py' + Qy'^2 = 0$ may be reduced to quadratures 1° when P and Q are both functions of y, or 2° when both are functions of x, or 3° when P is a function of x and Q is a function of y (integrating factor $1/y'$). In each case find the general expression for y in terms of quadratures. Integrate $y'' + 2y'\cot x + 2y'^2\tan y = 0$.

8. Find and discuss the curves for which the radius of curvature is proportional to the radius r of the curve.

9. If the radius of curvature R is expressed as a function $R = R(s)$ of the arc s measured from some point, the equation $R = R(s)$ or $s = s(R)$ is called the *intrinsic equation* of the curve. To find the relation between x and y the second equation may be differentiated as $ds = s'(R)\,dR$, and this equation of the third order may be solved. Show that if the origin be taken on the curve at the point $s = 0$ and if the x-axis be tangent to the curve, the equations

$$x = \int_0^s \cos\left[\int_0^s \frac{ds}{R}\right] ds, \qquad y = \int_0^s \sin\left[\int_0^s \frac{ds}{R}\right] ds$$

express the curve parametrically. Find the curves whose intrinsic equations are

$$(\alpha)\ R = a, \qquad (\beta)\ aR = s^2 + a^2, \qquad (\gamma)\ R^2 + s^2 = 16a^2.$$

10. Given $F = y^{(n)} + X_1 y^{(n-1)} + X_2 y^{(n-2)} + \cdots + X_{n-1}y' + X_n y = 0$. Show that if μ, a function of x alone, is an integrating factor of the equation, then

$$\Phi = \mu^{(n)} - (X_1\mu)^{(n-1)} + (X_2\mu)^{(n-2)} - \cdots + (-1)^{n-1}(X_{n-1}\mu)' + (-1)^n X_n\mu = 0$$

is the equation satisfied by μ. Collect the coefficient of μ to show that the condition that the given equation be exact is the condition that this coefficient vanish. The equation $\Phi = 0$ is called the *adjoint* of the given equation $F = 0$. Any integral μ of the adjoint equation is an integrating factor of the original equation. Moreover note that

$$\int \mu F dx = \mu y^{(n-1)} + (\mu X_1 - \mu')\, y^{(n-2)} + \cdots + (-1)^n \int y\Phi dx,$$

or
$$d[\mu F - (-1)^n y\Phi] = d\,[\mu y^{(n-1)} + (\mu X_1 - \mu')\, y^{(n-2)} + \cdots] = d\Omega.$$

Hence if μF is an exact differential, so is $y\Phi$. In other words, any solution y of the original equation is an integrating factor for the adjoint equation.

104. Linear differential equations. The equations

$$\begin{aligned} X_0 D^n y + X_1 D^{n-1} y + \cdots + X_{n-1} D y + X_n y &= R(x), \\ X_0 D^n y + X_1 D^{n-1} y + \cdots + X_{n-1} D y + X_n y &= 0 \end{aligned} \tag{20}$$

are linear differential equations of the nth order; the first is called the *complete equation* and the second the *reduced equation.* If $y_1, y_2, y_3, \cdots$ are any solutions of the reduced equation, and $C_1, C_2, C_3, \cdots$ are any constants, then $y = C_1 y_1 + C_2 y_2 + C_3 y_3 + \cdots$ is also a solution of the reduced equation. This follows at once from the linearity of the reduced equation and is proved by direct substitution. Furthermore if I is any solution of the complete equation, then $y + I$ is also a solution of the complete equation (cf. § 96).

As the equations (20) are of the nth order, they will determine $y^{(n)}$ and, by differentiation, all higher derivatives in terms of the values of $x, y, y', \cdots, y^{(n-1)}$. Hence if the values of the n quantities $y_0, y_0', \cdots, y_0^{(n-1)}$ which correspond to the value $x = x_0$ be given, all the higher derivatives are determined (§§ 87–88). Hence there are n and no more than n arbitrary conditions that may be imposed as initial conditions. A solution of the equations (20) which contains n distinct arbitrary constants is called the general solution. By distinct is meant that the constants can actually be determined to suit the n initial conditions.

If $y_1, y_2, \cdots, y_n$ are n solutions of the reduced equation, and

$$\begin{aligned} y &= C_1y_1 + C_2y_2 + \cdots + C_ny_n, \\ y' &= C_1y_1' + C_2y_2' + \cdots + C_ny_n', \\ &\cdots\cdots\cdots \\ y^{(n-1)} &= C_1y_1^{(n-1)} + C_2y_2^{(n-1)} + \cdots + C_ny_n^{(n-1)}, \end{aligned} \tag{21}$$

then y is a solution and $y', \cdots, y^{(n-1)}$ are its first $n-1$ derivatives. If x_0 be substituted on the right and the assumed corresponding initial values $y_0, y_0', \cdots, y_0^{(n-1)}$ be substituted on the left, the above n equations become linear equations in the n unknowns $C_1, C_2, \cdots, C_n$; and if they are to be soluble for the C's, the condition

$$W(y_1, y_2, \cdots, y_n) = \begin{vmatrix} y_1 & y_2 & \cdots & y_n \\ y_1' & y_2' & \cdots & y_n' \\ \cdot & \cdot & \cdot & \cdot \\ y_1^{(n-1)} & y_2^{(n-1)} & \cdots & y_n^{(n-1)} \end{vmatrix} \neq 0 \tag{22}$$

must hold for every value of $x = x_0$. Conversely if the condition does hold, the equations will be soluble for the C's.

The determinant $W(y_1, y_2, \cdots, y_n)$ is called the *Wronskian* of the n functions $y_1, y_2, \cdots, y_n$. The result may be stated as: If n functions $y_1, y_2, \cdots, y_n$ which are solutions of the reduced equation, and of which the Wronskian does not vanish, can be found, the general solution of the reduced equation can be written down. In general no solution of the equation can be found, whether by a definite process or by inspection; but in the rare instances in which the n solutions can be seen by inspection the problem of the solution of the reduced equation is completed. Frequently one solution may be found by inspection, and it is therefore important to see how much this contributes toward effecting the solution.

If y_1 is a solution of the reduced equation, make the substitution $y = y_1z$. The derivatives of y may be obtained by Leibniz's Theorem (§ 8). As the formula is linear in the derivatives of z, it follows that the result of the substitution will leave the equation linear in the new variable z. Moreover, to collect the coefficient of z itself, it is necessary to take only the first term $y_1^{(k)}z$ in the expansions for the derivative $y^{(k)}$. Hence

$$(X_0y_1^{(n)} + X_1y_1^{(n-1)} + \cdots + X_{n-1}y_1' + X_ny_1)z = 0$$

is the coefficient of z and vanishes by the assumption that y_1 is a solution of the reduced equation. Then the equation for z is

$$P_0z^{(n)} + P_1z^{(n-1)} + \cdots + P_{n-2}z'' + P_{n-1}z' = 0; \tag{23}$$

and if z' be taken as the variable, the equation is of the order $n-1$. It therefore appears that *the knowledge of a solution y_1 reduces the order of the equation by one.*

Now if $y_2, y_3, \cdots, y_p$ were other solutions, the derived ratios

$$z_1' = \left(\frac{y_2}{y_1}\right)', \qquad z_2' = \left(\frac{y_3}{y_1}\right)', \qquad \cdots, \qquad z_{p-1}' = \left(\frac{y_p}{y_1}\right)' \tag{23'}$$

would be solutions of the equation in z'; for by substitution,

$$y = y_1 z_1 = y_2, \qquad y = y_1 z_2 = y_3, \qquad \cdots, \qquad y = y_1 z_{p-1} = y_p$$

are all solutions of the equation in y. Moreover, if there were a linear relation $C_1 z_1' + C_2 z_2' + \cdots + C_{p-1} z_{p-1}' = 0$ connecting the solutions z_i', an integration would give a linear relation

$$C_1 y_2 + C_2 y_3 + \cdots + C_{p-1} y_n + C_p y_1 = 0$$

connecting the p solutions y_i. Hence if there is no linear relation (of which the coefficients are not all zero) connecting the p solutions y_i of the original equation, there can be none connecting the $p-1$ solutions z_i' of the transformed equation. Hence *a knowledge of p solutions of the original reduced equation gives a new reduced equation of which $p-1$ solutions are known.* And the process of substitution may be continued to reduce the order further until the order $n-p$ is reached.

As an example consider the equation of the third order

$$(1-x)y''' + (x^2-1)y'' - x^2 y + xy = 0.$$

Here a simple trial shows that x and e^x are two solutions. Substitute

$$y = e^x z, \qquad y' = e^x(z+z'), \qquad y'' = e^x(2z'+z''), \qquad y''' = e^x(3z''+z''').$$

Then $\qquad (1-x)z''' + (x^2-3x+2)z'' + (x^2-3x+1)z' = 0$

is of the second order in z'. A known solution is the derived ratio $(x/e^x)'$.

$$z' = (xe^{-x})' = e^{-x}(1-x). \quad \text{Let } z' = e^{-x}(1-x)w.$$

From this, z'' and z''' may be found and the equation takes the form

$$(1-x)w'' + (1+x)(x-2)w' = 0 \quad \text{or} \quad \frac{dw'}{w'} = x\,dx - \frac{2}{x-1}dx.$$

This is a linear equation of the first order and may be solved.

$$\log w' = \tfrac{1}{2}x^2 - 2\log(x-1) + C \quad \text{or} \quad w' = C_1 e^{\frac{1}{2}x^2}(x-1)^{-2}.$$

Hence
$$w = C_1 \int e^{\frac{1}{2}x^2}(x-1)^{-2}dx + C_2,$$

$$z' = \left(\frac{x}{e^x}\right)' w = C_1 \left(\frac{x}{e^x}\right)' \int e^{\frac{1}{2}x^2}(x-1)^{-2}dx + C_2\left(\frac{x}{e^x}\right)',$$

$$z = C_1 \int \left(\frac{x}{e^x}\right)' \int e^{\frac{1}{2}x^2}(x-1)^{-2}(dx)^2 + C_2\frac{x}{e^x} + C_3,$$

$$y = e^x z = C_1 e^x \int \left(\frac{x}{e^x}\right)' \int e^{\frac{1}{2}x^2}(x-1)^{-2}(dx)^2 + C_2 x + C_3 e^x.$$

The value for y is thus obtained in terms of quadratures. It may be shown that in case the equation is of the nth degree with p known solutions, the final result will call for $p(n-p)$ quadratures.

105. If the general solution $y = C_1y_1 + C_2y_2 + \cdots + C_ny_n$ of the reduced equation has been found (called the *complementary function* for the complete equation), the general solution of the complete equation may always be obtained in terms of quadratures by the important and far-reaching *method of the variation of constants* due to Lagrange. The question is: Cannot functions of x be found so that the expression

$$y = C_1(x)y_1 + C_2(x)y_2 + \cdots + C_n(x)y_n \tag{24}$$

shall be the solution of the complete equation? As there are n of these functions to be determined, it should be possible to impose $n-1$ conditions upon them and still find the functions.

Differentiate y on the supposition that the C's are variable.

$$y' = C_1y_1' + C_2y_2' + \cdots + C_ny_n' + y_1C_1' + y_2C_2' + \cdots + y_nC_n'.$$

As one of the conditions on the C's suppose that

$$y_1C_1' + y_2C_2' + \cdots + y_nC_n' = 0.$$

Differentiate again and impose the new condition

$$y_1'C_1' + y_2'C_2' + \cdots + y_n'C_n' = 0,$$

so that
$$y'' = C_1y_1'' + C_2y_2'' + \cdots + C_ny_n''.$$

The differentiation may be continued to the $(n-1)$st condition

$$y_1^{(n-2)}C_1' + y_2^{(n-2)}C_2' + \cdots + y_n^{(n-2)}C_n' = 0,$$

and
$$y^{(n-1)} = C_1y_1^{(n-1)} + C_2y_2^{(n-1)} + \cdots + C_ny_n^{(n-1)}.$$

Then
$$y^{(n)} = C_1y_1^{(n)} + C_2y_2^{(n)} + \cdots + C_ny_n^{(n)} + y_1^{(n-1)}C_1' + y_2^{(n-1)}C_2' + \cdots + y_n^{(n-1)}C_n'.$$

Now if the expressions thus found for $y, y', y'', \cdots, y^{(n-1)}, y^{(n)}$ be substituted in the complete equation, and it be remembered that $y_1, y_2, \cdots, y_n$ are solutions of the reduced equation and hence give 0 when substituted in the left-hand side of the equation, the result is

$$y_1^{(n-1)}C_1' + y_2^{(n-1)}C_2' + \cdots + y_n^{(n-1)}C_n' = R.$$

Hence, in all, there are n linear equations

$$\begin{aligned} y_1C_1' &+ y_2C_2' + \cdots + y_nC_n' = 0, \\ y_1'C_1' &+ y_2'C_2' + \cdots + y_n'C_n' = 0, \\ &\cdots\cdots\cdots \\ y_1^{(n-2)}C_1' &+ y_2^{(n-2)}C_2' + \cdots + y_n^{(n-2)}C_n' = 0, \\ y_1^{(n-1)}C_1' &+ y_2^{(n-1)}C_2' + \cdots + y_n^{(n-1)}C_n' = R. \end{aligned} \tag{25}$$

connecting the derivatives of the C's; and these may actually be solved for those derivatives which will then be expressed in terms of x. The C's may then be found by quadrature.

As an example consider the equation with constant coefficients

$$(D^3 + D)y = \sec x \quad \text{with} \quad y = C_1 + C_2 \cos x + C_3 \sin x$$

as the solution of the reduced equation. Here the solutions y_1, y_2, y_3 may be taken as 1, $\cos x$, $\sin x$ respectively. The conditions on the derivatives of the C's become by direct substitution in (25)

$$C_1' + \cos x C_2' + \sin x C_3' = 0, \quad -\sin x C_2' + \cos x C_3' = 0, \quad -\cos x C_2' - \sin x C_3' = \sec x.$$

Hence $\qquad C_1' = \sec x, \qquad C_2' = -1, \qquad C_3' = -\tan x$

and $\quad C_1 = \log\tan(\tfrac{1}{2}x + \tfrac{1}{4}\pi) + c_1, \qquad C_2 = -x + c_2, \qquad C_3 = \log\cos x + c_3.$

Hence $\quad y = c_1 + \log\tan(\tfrac{1}{2}x + \tfrac{1}{4}\pi) + (c_2 - x)\cos x + (c_3 + \log\cos x)\sin x$

is the general solution of the complete equation. This result could not be obtained by any of the real short methods of §§ 96–97. It could be obtained by the general method of § 95, but with little if any advantage over the method of variation of constants here given. The present method is equally available for equations with variable coefficients.

106. *Linear equations of the second order* are especially frequent in practical problems. In a number of cases the solution may be found. Thus 1° when the coefficients are constant or may be made constant by a change of variable as in Ex. 7, p. 222, the general solution of the reduced equation may be written down at once. The solution of the complete equation may then be found by obtaining a particular integral I by the methods of §§ 95–97 or by the application of the method of variation of constants. And 2° when the equation is exact, the solution may be had by integrating the linear equation (19) of § 103 of the first order by the ordinary methods. And 3° when one solution of the reduced equation is known (§ 104), the reduced equation may be completely solved and the complete equation may then be solved by the method of variation of constants, or the complete equation may be solved directly by Ex. 6 below.

Otherwise, write the differential equation in the form

$$\frac{d^2y}{dx^2} + P\frac{dy}{dx} + Q = R. \tag{26}$$

The substitution $y = uz$ gives the new equation

$$\frac{d^2z}{dx^2} + \left(\frac{2}{u}\frac{du}{dx} + P\right)\frac{dz}{dx} + \frac{1}{u}(u'' + Pu' + Qu)z = \frac{R}{u}. \tag{26'}$$

If u be determined so that the coefficient of z' vanishes, then

$$u = e^{-\frac{1}{2}\int Pdx} \quad \text{and} \quad \frac{d^2z}{dx^2} + \left(Q - \frac{1}{2}\frac{dP}{dx} - \frac{1}{4}P^2\right)z = Re^{\frac{1}{2}\int Pdx}. \tag{27}$$

Now 4° if $Q - \frac{1}{2}P' - \frac{1}{4}P^2$ is constant, the new reduced equation in (27) may be integrated; and 5° if it is k/x^2, the equation may also be integrated by the method of Ex. 7, p. 222. The integral of the complete equation may then be found. (In other cases this method may be useful in that the equation is reduced to a simpler form where solutions of the reduced equation are more evident.)

Again, suppose that the independent variable is changed to z. Then

$$\frac{d^2y}{dz^2} + \frac{z'' + Pz'}{z'^2}\frac{dy}{dz} + \frac{Q}{z'^2}y = \frac{R}{z'^2}. \tag{28}$$

Now 6° if $z'^2 = \pm Q$ will make $z'' + Pz' = kz'^2$, so that the coefficient of dy/dz becomes a constant k, the equation is integrable. (Trying if $z'^2 = \pm Qz^2$ will make $z'' + Pz' = kz'^2/z$ is needless because nothing in addition to 6° is thereby obtained. It may happen that if z be determined so as to make $z'' + Pz' = 0$, the equation will be so far simplified that a solution of the reduced equation becomes evident.)

Consider the example $\frac{d^2y}{dx^2} + \frac{2}{x}\frac{dy}{dx} + \frac{a^2}{x^4}y = 0$. Here no solution is apparent. Hence compute $Q - \frac{1}{2}P' - \frac{1}{4}P^2$. This is a^2/x^4 and is neither constant nor proportional to $1/x^2$. Hence the methods 4° and 5° will not work. From $z'^2 = Q = a^2/x^4$ or $z' = a/x^2$, it appears that $z'' + Pz' = 0$, and 6° works; the new equation is

$$\frac{d^2y}{dz^2} + y = 0 \quad \text{with} \quad z = -\frac{a}{x}.$$

The solution is therefore seen immediately to be

$$y = C_1 \cos z - C_2 \sin z \quad \text{or} \quad y = C_1 \cos(a/x) + C_2 \sin(a/x).$$

If there had been a right-hand member in the original equation, the solution could have been found by the method of variation of constants, or by some of the short methods for finding a particular solution if R had been of the proper form.

EXERCISES

1. If a relation $C_1y_1 + C_2y_2 + \cdots + C_ny_n = 0$, with constant coefficients not all 0, exists between n functions $y_1, y_2, \cdots, y_n$ of x for all values of x, the functions are by definition said to be *linearly dependent;* if no such relation exists, they are said to be *linearly independent.* Show that the nonvanishing of the Wronskian is a criterion for linear independence.

2. If the general solution $y = C_1y_1 + C_2y_2 + \cdots + C_ny_n$ is the same for

$$X_0y^{(n)} + X_1y^{(n-1)} + \cdots + X_ny = 0 \quad \text{and} \quad P_0y^{(n)} + P_1y^{(n-1)} + \cdots + P_ny = 0,$$

two linear equations of the nth order, show that y satisfies the equation

$$(X_1P_0 - X_0P_1)y^{(n-1)} + \cdots + (X_nP_0 - X_0P_n)y = 0$$

of the $(n-1)$st order; and hence infer, from the fact that y contains n arbitrary constants corresponding to n arbitrary initial conditions, the important theorem: If two linear equations of the nth order have the same general solution, the corresponding coefficients are proportional.

3. If $y_1, y_2, \cdots, y_n$ are n independent solutions of an equation of the nth order, show that the equation may be taken in the form $W(y_1, y_2, \cdots, y_n, y) = 0$.

4. Show that if, in any reduced equation, $X_{n-1} + xX_n = 0$ identically, then x is a solution. Find the condition that x^m be a solution; also that e^{mx} be a solution.

5. Find by inspection one or more independent solutions and integrate:

(α) $(1+x^2)y'' - 2xy' + 2y = 0$, ($\beta$) $xy'' + (1-x)y' - y = 0$,

(γ) $(ax - bx^2)y'' - ay' + 2by = 0$, ($\delta$) $\frac{1}{2}y'' + xy' - (x+2)y = 0$,

(ϵ) $\left(\log x + \frac{1}{x^4} - \frac{1}{x^2} + \frac{1}{x}\right)y''' + \left(\log x + \frac{1}{x^4} + \frac{1}{x^3} - \frac{1}{x^2}\right)y'' + \left(\frac{1}{x^2} - \frac{1}{x}\right)(y' - xy) = 0$,

(ζ) $y^{\text{iv}} - xy''' + xy' - y = 0$, ($\eta$) $(4x^2 - x + 1)y''' + 8x^2y'' - 4xy' - 8y = 0$.

6. If y_1 is a known solution of the equation $y'' + Py' + Qy = R$ of the second order, show that the general solution may be written as

$$y = C_1y_1 + C_2y_1\int e^{-\int Pdx}\frac{dx}{y_1^2} + y_1\int \frac{1}{y_1^2}e^{-\int Pdx}\int y_1 e^{\int Pdx}R\,(dx)^2.$$

7. Integrate:

(α) $xy'' - (2x+1)y' + (x+1)y = x^2 - x - 1$,

(β) $y'' - x^2y' + xy = x$, (γ) $xy'' + (1-x)y' - y = e^x$,

(δ) $y'' - xy' + (x-1)y = R$, (ϵ) $y''\sin^2 x + y'\sin x\cos x - y = x - \sin x$.

8. After writing down the integral of the reduced equation by inspection, apply the method of the variation of constants to these equations:

(α) $(D^2+1)y = \tan x$, (β) $(D^2+1)y = \sec^2 x$, (γ) $(D-1)^2y = e^x(1-x)^{-2}$,

(δ) $(1-x)y'' + xy' - y = (1-x)^2$, ($\epsilon$) $(1 - 2x + x^2)(y''' - 1) - x^2y'' + 2xy' - y = 1$.

9. Integrate the following equations of the second order:

(α) $4x^2y'' + 4x^3y' + (x^2+1)^2y = 0$, ($\beta$) $y'' - 2y'\tan x - (a^2+1)y = 0$,

(γ) $xy'' + 2y' - xy = 2e^x$, (δ) $y''\sin x + 2y'\cos x + 3y\sin x = e^x$,

(ϵ) $y'' + y'\tan x + y\cos^2 x = 0$, ($\zeta$) $(1-x^2)y'' - xy' + 4y = 0$,

(η) $y'' + (2e^x - 1)y' + e^{2x}y = e^{4x}$, ($\theta$) $x^6y'' + 3x^5y' + y = x^{-2}$.

10. Show that if $X_0y'' + X_1y' + X_2y = R$ may be written in factors as

$$(X_0D^2 + X_1D + X_2)y = (p_1D + q_1)(p_2D + q_2)y = R,$$

where the factors are not commutative inasmuch as the differentiation in one factor is applied to the variable coefficients of the succeeding factor as well as to D, then the solution is obtainable in terms of quadratures. Show that

$$q_1p_2 + p_1p_2' + p_1q_2 = X_1 \quad \text{and} \quad q_1q_2 + p_1q_2' = X_2.$$

In this manner integrate the following equations, choosing p_1 and p_2 as factors of X_0 and determining q_1 and q_2 by inspection or by assuming them in some form and applying the method of undetermined coefficients:

(α) $xy'' + (1-x)y' - y = e^x$, (β) $3x^2y'' + (2 - 6x^2)y' - 4 = 0$,

(γ) $3x^2y'' + (2 + 6x - 6x^2)y' - 4y = 0$, ($\delta$) $(x^2-1)y'' - (3x+1)y' - x(x-1)y = 0$,

(ϵ) $axy'' + (3a + bx)y' + 3by = 0$, ($\zeta$) $xy'' - 2x(1+x)y' + 2(1+x)y = x^3$.

11. Integrate these equations in any manner:

(α) $y'' - \frac{1}{\sqrt{x}}y' + \frac{x + \sqrt{x} - 8}{4x^2}y = 0$, ($\beta$) $y'' - \frac{2}{x}y' + \left(a^2 + \frac{2}{x^2}\right)y = 0$,

(γ) $y'' + y'\tan x + y\cos^2 x = 0$, ($\delta$) $y''-2\left(n-\frac{a}{x}\right)y'+\left(n^2-2\frac{na}{x}\right)y=e^{nx}$,

(ϵ) $(1-x^2)y'' - xy' - c^2y = 0$, ($\zeta$) $(a^2-x^2)y'' - 8xy' - 12y = 0$,

(η) $y'' + \frac{1}{x^2\log x}y = e^x\left(\frac{2}{x} + \log x\right)$, ($\theta$) $y'' - \frac{9-4x}{3-x}y' + \frac{6-3x}{3-x}y = 0$,

(ι) $y'' + 2x^{-1}y' - n^2y = 0$, ($\kappa$) $y'' - 4xy' + (4x^2-3)y = e^{x^2}$,

(λ) $y'' + 2ny'\cot nx + (m^2-n^2)y = 0$, ($\mu$) $y'' + 2(x^{-1}+Bx^{-2})y' + Ax^{-4}y = 0$.

12. If y_1 and y_2 are solutions of $y'' + Py' + R = 0$, show by eliminating Q and integrating that

$$y_1y_2' - y_2y_1' = Ce^{-\int Pdx}.$$

What if $C = 0$? If $C \neq 0$, note that y_1 and y_1' cannot vanish together; and if $y_1(a) = y_1(b) = 0$, use the relation $(y_2y_1')_a : (y_2y_1')_b = k > 0$ to show that as y_{1a}' and y_{1b}' have opposite signs, y_{2a} and y_{2b} have opposite signs and hence $y_2(\xi) = 0$ where $a < \xi < b$. Hence the theorem: Between any two roots of a solution of an equation of the second order there is one root of every solution independent of the given solution. What conditions of continuity for y and y' are tacitly assumed here?

107. The cylinder functions. Suppose that $C_n(x)$ is a function of x which is different for different values of n and which satisfies the two equations

$$C_{n-1}(x) - C_{n+1}(x) = 2\frac{d}{dx}C_n(x), \quad C_{n-1}(x) + C_{n+1}(x) = \frac{2n}{x}C_n(x). \tag{29}$$

Such a function is called a *cylinder function* and the index n is called the *order* of the function and may have any real value. The two equations are supposed to hold for all values of n and for all values of x. They do not completely determine the functions but from them follow the chief rules of operation with the functions. For instance, by addition and subtraction,

$$C_n'(x) = C_{n-1}(x) - \frac{n}{x}C_n(x) = \frac{n}{x}C_n(x) - C_{n+1}(x). \tag{30}$$

Other relations which are easily deduced are

$$D_x[x^nC_n(ax)] = ax^nC_{n-1}(ax), \quad D_x[x^{-n}C_n(ax)] = -ax^{-n}C_{n+1}(x), \tag{31}$$

$$D_x\left[x^{\frac{n}{2}}C_n(\sqrt{ax})\right] = \tfrac{1}{2}\sqrt{a}x^{\frac{n-1}{2}}C_{n-1}(\sqrt{ax}), \tag{32}$$

$$C_0'(x) = -C_1(x), \quad C_{-n}(x) = (-1)^nC_n(x), \quad n \text{ integral}, \tag{33}$$

$$C_n(x)K_n'(x) - C_n'(x)K_n(x) = C_{n+1}(x)K_n(x) - C_n(x)K_{n+1}(x) = \frac{A}{x}, \tag{34}$$

where C and K denote any two cylinder functions.

The proof of these relations is simple, but will be given to show the use of (29). In the first case differentiate directly and substitute from (29).

$$D_x[x^nC_n(\alpha x)] = x^n\left[\alpha D_{\alpha x}C_n(\alpha x) + \frac{n}{x}C_n(\alpha x)\right]$$
$$= x^n\left[\alpha C_{n-1}(\alpha x) - \alpha\frac{n}{\alpha x}C_n(\alpha x) + \frac{n}{x}C_n(\alpha x)\right].$$

The second of (31) is proved similarly. For (32), differentiate.

$$D_x[x^{\frac{n}{2}}C_n(\sqrt{\alpha x})] = \frac{1}{2}nx^{\frac{n}{2}-1}C_n(\sqrt{\alpha x}) + x^{\frac{n}{2}}\frac{1}{2}\sqrt{\frac{\alpha}{x}}D_{\sqrt{\alpha x}}C_n(\sqrt{\alpha x})$$

$$= \frac{1}{2}\sqrt{\alpha}x^{\frac{n-1}{2}}\left[\frac{n}{\sqrt{\alpha x}}C_n(\sqrt{\alpha x}) + C_{n-1}(\sqrt{\alpha x}) - \frac{n}{\sqrt{\alpha x}}C_n(\sqrt{\alpha x})\right].$$

Next (33) is obtained 1° by substituting 0 for n in both equations (29).

$$C_{-1}(x) - C_1(x) = 2\,C_0'(x), \quad C_{-1}(x) + C_1(x) = 0, \quad \text{hence} \quad C_0'(x) = -\,C_1(x);$$

and 2° by substituting successive values for n in the second of (29) written in the form $xC_{n-1} + xC_{n+1} = 2\,nC_n$. Then

$$xC_{-1} + xC_1 = 0, \qquad xC_{-2} + xC_0 = -\,2\,C_{-1}, \qquad xC_0 + xC_2 = 2\,C_1,$$
$$xC_{-3} + xC_{-1} = -\,4\,C_{-2}, \qquad xC_1 + xC_3 = 4\,C_2,$$
$$xC_{-4} + xC_{-2} = -\,6\,C_3, \qquad xC_2 + xC_4 = 6\,C_3,$$

and so on. The first gives $C_{-1} = -\,C_1$. Subtract the next two and use $C_{-1} + C_1 = 0$. Then $C_{-2} - C_2 = 0$ or $C_{-2} = (-1)^2C_2$. Add the next two and use the relations already found. Then $C_{-3} + C_3 = 0$ or $C_{-3} = (-1)^3C_3$. Subtract the next two, and so on. For the last of the relations, a very important one, note first that the two expressions become equivalent by virtue of (29); for

$$C_nK_n' - C_n'K_n = \frac{n}{x}C_nK_n - C_nK_{n+1} - \frac{n}{x}C_nK_n + C_{n+1}K_n.$$

Now $\dfrac{d}{dx}[x(C_{n+1}K_n - C_nK_{n+1})] = C_{n+1}K_n - C_nK_{n+1} + xK_n\left(C_n - \dfrac{n+1}{x}C_{n+1}\right)$

$$+\,xC_{n+1}\left(\frac{n}{x}K_n - K_{n+1}\right) - xK_{n+1}\left(\frac{n}{x}C_n - C_{n+1}\right)$$
$$-\,xC_n\left(K_n - \frac{n+1}{x}K_{n+1}\right) = 0.$$

Hence $x(C_{n+1}K_n - C_nK_{n+1}) = \text{const.} = A$, and the relation is proved.

The cylinder functions of a given order n satisfy a linear differential equation of the second order. This may be obtained by differentiating the first of (29) and combining with (30).

$$2\,C_n'' = C_{n-1}' - C_{n+1}' = \frac{n-1}{x}C_{n-1} - 2\,C_n + \frac{n+1}{x}C_{n+1}$$
$$= \frac{n}{x}(C_{n-1} + C_{n+1}) - \frac{1}{x}(C_{n-1} - C_{n+1}) - 2\,C_n.$$

Hence
$$\frac{d^2y}{dx^2} + \frac{1}{x}\frac{dy}{dx} + \left(1 - \frac{n^2}{x^2}\right)y = 0, \qquad y = C_n(x). \tag{35}$$

This equation is known as *Bessel's equation;* the functions $C_n(x)$, which have been called cylinder functions, are often called *Bessel's functions.* From the equation it follows that any three functions of the same order n are connected by a linear relation and there are only two independent functions of any given order.

By a change of the independent variable, the Bessel equation may take on several other forms. The easiest way to find them is to operate directly with the relations (31), (32). Thus

$$D_x[x^{-n}C_n(x)] = -x^{-n}C_{n+1} = -x \cdot x^{-n-1}C_{n+1},$$
$$D_x^2[x^{-n}C_n(x)] = -x^{-n-1}C_{n+1} + x \cdot x^{-n-1}C_{n+2}$$
$$= -x^{-n-1}C_{n+1} + 2(n+1)x^{-n-1}C_{n+1} - x^{-n}C_n.$$

Hence $$\frac{d^2y}{dx^2} + \frac{(1+2n)}{x}\frac{dy}{dx} + y = 0, \qquad y = x^{-n}C_n(x). \tag{36}$$

Again $$\frac{d^2y}{dx^2} + \frac{(1-2n)}{x}\frac{dy}{dx} + y = 0, \qquad y = x^nC_n(x). \tag{37}$$

Also $$xy'' + (1+n)y' + y = 0, \qquad y = x^{-\frac{n}{2}}C_n(2\sqrt{x}). \tag{38}$$

And $$xy'' + (1-n)y' + y = 0, \qquad y = x^{\frac{n}{2}}C_n(2\sqrt{x}). \tag{39}$$

In all these differential equations it is well to restrict x to positive values inasmuch as, if n is not specialized, the powers of x, as x^n, x^{-n}, $x^{\frac{n}{2}}$, $x^{-\frac{n}{2}}$, are not always real.

108. The fact that n occurs only squared in (35) shows that both $C_n(x)$ and $C_{-n}(x)$ are solutions, so that if these functions are independent, the complete solution is $y = aC_n + bC_{-n}$. In like manner the equations (36), (37) form a pair which differ only in the sign of n. Hence if H_n and H_{-n} denote particular integrals of the first and second respectively, the complete integrals are respectively

$$y = aH_n + bH_{-n}x^{-2n} \quad \text{and} \quad y = aH_{-n} + bH_nx^{2n};$$

and similarly the respective integrals of (38), (39) are

$$y = aI_n + bI_{-n}x^{-n} \quad \text{and} \quad y = aI_{-n} + bI_nx^n,$$

where I_n and I_{-n} denote particular integrals of these two equations. It should be noted that these forms are the complete solutions only when the two integrals are independent. Note that

$$I_n(x) = x^{-\frac{1}{2}n}C_n(2\sqrt{x}), \qquad C_n(x) = (\tfrac{1}{2}x)^nI_n(\tfrac{1}{4}x^2). \tag{40}$$

As it has been seen that $C_n = (-1)^nC_{-n}$ when n is integral, it follows that in this case the above forms do not give the complete solution.

A particular solution of (38) may readily be obtained in series by the method of undetermined coefficients (§ 88). It is

$$I_n(x) = \sum_0^\infty a_ix^i, \qquad a_i = \frac{(-1)^i}{i!(n+1)(n+2)\cdots(n+i)}, \tag{41}$$

as is derived below. It should be noted that I_{-n} formed by changing the sign of n is meaningless when n is an integer, for the reason that from a certain point on, the coefficients a_i have zeros in the denominator. The determination of a series for the second independent solution when n is integral will be omitted. The solutions of (35), (36) corresponding to $I_n(x)$ are, by (40) and (41),

$$J_n(x) = \frac{x^n}{2^n}\sum_0^\infty \frac{(-1)^i x^{2i}}{2^{2i} i!\,(n+i)!} = \frac{x^n}{2^n n!} I_n(\tfrac{1}{4}x^2), \tag{42}$$

$$x^{-n}J_n(x) = \frac{1}{2^n n!} I_n(\tfrac{1}{4}x^2), \tag{42'}$$

where the factor $n!$ has been introduced in the denominator merely to conform to usage.* The chief cylinder function $C_n(x)$ is $J_n(x)$ and it always carries the name of Bessel.

To derive the series for $I_n(x)$ write

$$\begin{array}{r|l} 1 & I_n = a_0 + a_1x + a_2x^2 + \cdots + a_{k-1}x^{k-1} + \cdots, \\ (1+n) & I_n' = a_1 + 2a_2x + 3a_3x^2 + \cdots + (k-1)a_{k-1}x^{k-2} + \cdots, \\ x & I_n'' = 2a_2 + 3\cdot 2a_3x + \cdots + (k-1)(k-2)a_{k-1}x^{k-3} + \cdots, \\ \hline & 0 = [a_0 + a_1(n+1)] + x[a_1 + a_2 2(n+2)] + x^2[a_2 + a_3 3(n+3)] \\ & \qquad + \cdots + x^{k-1}[a_{k-1} + a_k k(n+k)] + \cdots. \end{array}$$

Hence $a_0 + a_1(n+1) = 0,\quad a_1 + a_2 2(n+2) = 0, \cdots,\quad a_{k-1} + a_k k(n+k) = 0,$

or

$$a_1 = -\frac{a_0}{n+1},\qquad a_2 = \frac{-a_1}{2(n+2)} = \frac{a_0}{2!\,(n+1)(n+2)}, \cdots,$$

$$a_k = \frac{(-1)^k a_0}{k!\,(n+1)\cdots(n+k)}.$$

If now the choice $a_0 = 1$ is made, the series for $I_n(x)$ is as given in (41).

The famous differential equation of the first order

$$xy' - ay + by^2 = cx^n, \tag{43}$$

known as *Riccati's equation*, may be integrated in terms of cylinder functions. Note that if $n = 0$ or $c = 0$, the variables are separable; and if $b = 0$, the equation is linear. As these cases are immediately integrable, assume $bcn \neq 0$. By a suitable change of variable, the equation takes the form

$$\xi\frac{d^2\eta}{d\xi^2} + \left(1 - \frac{a}{n}\right)\frac{d\eta}{d\xi} - bc\eta = 0,\qquad \xi = \frac{1}{n^2}x^n,\qquad y = \frac{n}{b}\frac{d\eta}{d\xi}\frac{\xi}{\eta}. \tag{43'}$$

A comparison of this with (39) shows that the solution is

$$\eta = AI_{-\frac{a}{n}}(-bc\xi) + BI_{\frac{a}{n}}(-bc\xi)\cdot(-bc\xi)^{\frac{a}{n}},$$

which in terms of Bessel functions J becomes, by (40),

$$\eta = \xi^{\frac{a}{2n}}\left[AJ_{\frac{a}{n}}(2\sqrt{-bc\xi}) + BJ_{-\frac{a}{n}}(2\sqrt{-bc\xi})\right].$$

* If n is not integral, both $n!$ and $(n+i)!$ must be replaced (§ 147) by $\Gamma(n+1)$ and $\Gamma(n+i+1)$.

The value of y may be found by substitution and use of (29).

$$y=\sqrt{-\frac{c}{b}}\,x^{\frac{n}{2}}\,\frac{J_{\frac{a}{n}-1}(2x^{\frac{n}{2}}\sqrt{-bc}/n)-AJ_{1-\frac{a}{n}}(2x^{\frac{n}{2}}\sqrt{-bc}/n)}{J_{\frac{a}{n}}(2x^{\frac{n}{2}}\sqrt{-bc}/n)+AJ_{-\frac{a}{n}}(2x^{\frac{n}{2}}\sqrt{-bc}/n)}, \qquad (44)$$

where A denotes the one arbitrary constant of integration.

It is noteworthy that the cylinder functions are sometimes expressible in terms of trigonometric functions. For when $n=\frac{1}{2}$ the equation (35) has the integrals

$$y=A\sin x+B\cos x \quad \text{and} \quad y=x^{\frac{1}{2}}[AC_{\frac{1}{2}}(x)+BC_{-\frac{1}{2}}(x)].$$

Hence it is permissible to write the relations

$$x^{\frac{1}{2}}C_{\frac{1}{2}}(x)=\sin x, \qquad x^{\frac{1}{2}}C_{-\frac{1}{2}}(x)=\cos x, \qquad (45)$$

where C is a suitably chosen cylinder function of order $\frac{1}{2}$. From these equations by application of (29) the cylinder functions of order $p+\frac{1}{2}$, where p is any integer, may be found.

Now if Riccati's equation is such that b and c have opposite signs and a/n is of the form $p+\frac{1}{2}$, the integral (44) can be expressed in terms of trigonometric functions by using the values of the functions $C_{p+\frac{1}{2}}$ just found in place of the J's. Moreover if b and c have the same sign, the trigonometric solution will still hold formally and may be converted into exponential or hyperbolic form. Thus Riccati's equation is integrable in terms of the elementary functions when $a/n=p+\frac{1}{2}$ no matter what the sign of bc is.

EXERCISES

1. Prove the following relations:

(α) $4C_n''=C_{n-2}-2C_n+C_{n+2}$, ($\beta$) $xC_n=2(n+1)C_{n+1}-xC_{n+2}$,
(γ) $2^3C_n'''=C_{n-3}-3C_{n-1}+3C_{n+1}-C_{n+3}$, generalize,
(δ) $xC_n=2(n+1)C_{n+1}-2(n+3)C_{n+3}+2(n+5)C_{n+5}-xC_{n+6}$.

2. Study the functions defined by the pair of relations

$$F_{n-1}(x)+F_{n+1}(x)=2\frac{d}{dx}F_n(x), \qquad F_{n-1}(x)-F_{n+1}(x)=\frac{2}{x}F_n(x)$$

especially to find results analogous to (30)–(35).

3. Use Ex. 12, p. 247, to obtain (34) and the corresponding relation in Ex. 2.

4. Show that the solution of (38) is $y=AI_n\int\frac{dx}{x^{n+1}I_n^2}+BI_n$.

5. Write out five terms in the expansions of I_0, I_1, $I_{-\frac{1}{2}}$, J_0, J_1.

6. Show from the expansion (42) that $\frac{1}{2}!\sqrt{\frac{2}{x}}J_{\frac{1}{2}}(x)=\frac{1}{x}\sin x$.

7. From (45), (29) obtain the following:

$$x^{\frac{1}{2}}C_{\frac{3}{2}}(x)=\frac{\sin x}{x}-\cos x, \qquad x^{\frac{1}{2}}C_{\frac{5}{2}}(x)=\left(\frac{3}{x^2}-1\right)\sin x-\frac{3}{x}\cos x,$$

$$x^{\frac{1}{2}}C_{-\frac{3}{2}}(x)=-\sin x-\frac{\cos x}{x}, \qquad x^{\frac{1}{2}}C_{-\frac{5}{2}}(x)=\frac{3}{x}\sin x+\left(\frac{3}{x^2}-1\right)\cos x.$$

8. Prove by integration by parts: $\int \frac{J_2 dx}{x^3} dx = \frac{J_3}{x^3} + 6\frac{J_4}{x^4} + 6\cdot 8\int \frac{J_5 dx}{x^5}$.

9. Suppose $C_n(x)$ and $K_n(x)$ so chosen that $A = 1$ in (34). Show that

$$y = AC_n(x) + BK_n(x) + L\left[K_n(x)\int \frac{C_n(x)}{x^3}dx - C_n(x)\int \frac{K_n(x)}{x^3}dx\right]$$

is the integral of the differential equation $x^2y'' + xy' + (x^2 - n^2)y = Lx^{-2}$.

10. Note that the solution of Riccati's equation has the form

$$y = \frac{f(x) + Ag(x)}{F(x) + AG(x)}, \quad \text{and show that} \quad \frac{dy}{dx} + P(x)y + Q(x)y^2 = R(x)$$

will be the form of the equation which has such an expression for its integral.

11. Integrate these equations in terms of cylinder functions and reduce the results whenever possible by means of Ex. 7:

(α) $xy' - 5y + y^2 + x^2 = 0$, (β) $xy' - 3y + y^2 = x^2$,
(γ) $y'' + ye^{2x} = 0$, (δ) $x^2y'' + nxy' + (b + cx^{2m})y = 0$.

12. Identify the functions of Ex. 2 with the cylinder functions of ix.

13. Let $(x^2 - 1)P_n' = (n + 1)(P_{n+1} - xP_n)$, $P_{n+1}' = xP_n' + (n + 1)P_n$ (46)

be taken as defining the *Legendre functions* $P_n(x)$ of order n. Prove

(α) $(x^2 - 1)P_n' = n(xP_n - P_{n-1})$, (β) $(2n + 1)xP_n = (n + 1)P_{n+1} + nP_{n-1}$,
(γ) $(2n + 1)P_n = P_{n+1}' - P_{n-1}'$, (δ) $(1 - x^2)P_n'' - 2xP_n' + n(n + 1)P_n = 0$.

14. Show that $P_nQ_n' - P_n'Q_n = \frac{A}{x^2 - 1}$ and $P_nQ_{n+1} - P_{n+1}Q_n = \frac{A}{n + 1}$,

where P and Q are any two Legendre functions. Express the general solution of the differential equation of Ex. 13 (δ) analogously to Ex. 4.

15. Let $u = x^2 - 1$ and let D denote differentiation by x. Show

$$D^{n+1}u^{n+1} = D^{n+1}(uu^n) = uD^{n+1}u^n + 2(n + 1)xD^nu^n + n(n + 1)D^{n-1}u^n,$$
$$D^{n+1}u^{n+1} = D^nDu^{n+1} = 2(n + 1)D^n(xu^n) = 2(n + 1)xD^nu^n + 2n(n + 1)D^{n-1}u^n.$$

Hence show that the derivative of the second equation and the eliminant of $D^{n-1}u^n$ between the two equations give two equations which reduce to (46) if

$$P_n(x) = \frac{1}{2^n \cdot n!}\frac{d^n}{dx^n}(x^2 - 1)^n.$$ {When n is integral these are *Legendre's polynomials*.

16. Determine the solutions of Ex. 13 (δ) in series for the initial conditions

(α) $P_n(0) = 1$, $P_n'(0) = 0$, (β) $P_n(0) = 0$, $P_n'(0) = 1$.

17. Take $P_0 = 1$ and $P_1 = x$. Show that these are solutions of (46) and compute P_2, P_3, P_4 from Ex. 13 (β). If $x = \cos\theta$, show

$P_2 = \frac{3}{4}\cos 2\theta + \frac{1}{4}$, $P_3 = \frac{5}{8}\cos 3\theta + \frac{3}{8}\cos\theta$, $P_4 = \frac{35}{64}\cos 4\theta + \frac{20}{64}\cos 2\theta + \frac{9}{64}$.

18. Write Ex. 13 (δ) as $\frac{d}{dx}[(1 - x^2)P_n'] + n(n + 1)P_n = 0$ and show

$$[m(m + 1) - n(n + 1)]\int_{-1}^{+1} P_nP_m dx = \int_{-1}^{+1}\left[P_m\frac{d(1 - x^2)P_n'}{dx} - P_n\frac{d(1 - x^2)P_m'}{dx}\right]dx.$$

Integrate by parts, assume the functions and their derivatives are finite, and show

$$\int_{-1}^{+1} P_n P_m dx = 0, \quad \text{if} \quad n \neq m.$$

19. By successive integration by parts and by reduction formulas show

$$\int_{-1}^{+1} P_n^2 dx = \frac{1}{2^{2n}(n!)^2}\int_{-1}^{+1}\frac{d^n(x^2-1)^n}{dx^n}\cdot\frac{d^n(x^2-1)^n}{dx^n}dx = \frac{(-1)^n}{2^n\cdot n!}\int_{-1}^{+1}(x^2-1)^n dx$$

and

$$\int_{-1}^{+1} P_n^2 dx = \frac{2}{2n+1}, \qquad n \text{ integral}.$$

20. Show $\displaystyle\int_{-1}^{+1} x^m P_n dx = \int_{-1}^{+1} x^m \frac{d^n(x^2-1)^n}{dx^n} = 0, \qquad \text{if } m < n.$

Determine the value of the integral when $m = n$. Cannot the results of Exs. 18, 19 for m and n integral be obtained simply from these results?

21. Consider (38) and its solution $I_0 = 1 - x + \frac{x^2}{2!^2} - \frac{x^3}{3!^2} + \frac{x^4}{4!^2} - \cdots$ when $n = 0$. Assume a solution of the form $y = I_0 v + w$ so that

$$x\frac{d^2w}{dx^2} + \frac{dw}{dx} + w + 2x\frac{dI_0}{dx}\frac{dv}{dx} = 0, \quad \text{if} \quad x\frac{d^2v}{dx^2} + \frac{dv}{dx} = 0,$$

is the equation for w if v satisfies the equation $xv'' + v' = 0$. Show

$$v = A + B\log x, \qquad xw'' + w' + w = 2B - \frac{2Bx}{2!} + \frac{2Bx^2}{2!3!} - \frac{2Bx^3}{3!4!} + \cdots.$$

By assuming $w = a_1 x + a_2 x^2 + \cdots$, determine the a's and hence obtain

$$w = 2B\left[x - \frac{x^2}{2!^2}\left(1+\frac{1}{2}\right) + \frac{x^3}{3!^2}\left(1+\frac{1}{2}+\frac{1}{3}\right) - \frac{x^4}{4!^2}\left(1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}\right) + \cdots\right];$$

and $(A + B\log x)I_0 + w$ is then the complete solution containing two constants. As AI_0 is one solution, $B\log x\cdot I_0 + w$ is another. From this second solution for $n = 0$, the second solution for any integral value of n may be obtained by differentiation; the work, however, is long and the result is somewhat complicated.

CHAPTER X

DIFFERENTIAL EQUATIONS IN MORE THAN TWO VARIABLES

109. Total differential equations. An equation of the form

$$P(x, y, z)\,dx + Q(x, y, z)\,dy + R(x, y, z)\,dz = 0, \tag{1}$$

involving the differentials of three variables is called a *total differential equation.* A similar equation in any number of variables would also be called total; but the discussion here will be restricted to the case of three. If definite values be assigned to x, y, z, say a, b, c, the equation becomes

$$A\,dx + B\,dy + C\,dz = A(x-a) + B(y-b) + C(z-c) = 0, \tag{2}$$

where x, y, z are supposed to be restricted to values near a, b, c, and represents a small portion of a plane passing through (a, b, c). From the analogy to the lineal element (§ 85), such a portion of a plane may be called a *planar element.* The differential equation therefore represents an infinite number of planar elements, one passing through each point of space.

Now any family of surfaces $F(x, y, z) = C$ also represents an infinity of planar elements, namely, the portions of the tangent planes at every point of all the surfaces in the neighborhood of their respective points of tangency. In fact

$$dF = F'_x dx + F'_y dy + F'_z dz = 0 \tag{3}$$

is an equation similar to (1). If the planar elements represented by (1) and (3) are to be the same, the equations cannot differ by more than a factor $\mu(x, y, z)$. Hence

$$F'_x = \mu P, \qquad F'_y = \mu Q, \qquad F'_z = \mu R.$$

If a function $F(x, y, z) = C$ can be found which satisfies these conditions, it is said to be the integral of (1), and the factor $\mu(x, y, z)$ by which the equations (1) and (3) differ is called an *integrating factor* of (1). Compare § 91.

It may happen that $\mu = 1$ and that (1) is thus an *exact* differential. In this case the conditions

$$P'_y = Q'_x, \qquad Q'_z = R'_y, \qquad R'_x = P'_z, \tag{4}$$

which arise from $F''_{xy} = F''_{yx}$, $F''_{yz} = F''_{zy}$, $F''_{zx} = F''_{xz}$, must be satisfied. Moreover if these conditions are satisfied, the equation (1) will be an exact equation and the integral is given by

$$F(x, y, z) = \int_{x_0}^{x} P(x, y, z)\, dx + \int_{y_0}^{y} Q(x_0, y, z)\, dy + \int R(x_0, y_0, z)\, dz = C,$$

where x_0, y_0, z_0 may be chosen so as to render the integration as simple as possible. The proof of this is so similar to that given in the case of two variables (§ 92) as to be omitted. In many cases which arise in practice the equation, though not exact, may be made so by an obvious integrating factor.

As an example take $zxdy - yzdx + x^2dz = 0$. Here the conditions (4) are not fulfilled but the integrating factor $1/x^2z$ is suggested. Then

$$\frac{xdy - ydx}{x^2} + \frac{dz}{z} = d\left(\frac{y}{x} + \log z\right)$$

is at once perceived to be an exact differential and the integral is $y/x + \log z = C$. It appears therefore that in this simple case neither the renewed application of the conditions (4) nor the general formula for the integral was necessary. It often happens that both the integrating factor and the integral can be recognized at once as above.

If the equation does not suggest an integrating factor, the question arises, Is there any integrating factor? In the case of two variables (§ 94) there always was an integrating factor. In the case of three variables there may be none. For

$$\begin{array}{l|l}
F''_{xy} = P\dfrac{\partial \mu}{\partial y} + \mu\dfrac{\partial P}{\partial y} = F''_{yx} = Q\dfrac{\partial \mu}{\partial x} + \mu\dfrac{\partial Q}{\partial x}, & R, \\
F''_{yz} = Q\dfrac{\partial \mu}{\partial z} + \mu\dfrac{\partial Q}{\partial z} = F''_{zy} = R\dfrac{\partial \mu}{\partial y} + \mu\dfrac{\partial R}{\partial y}, & P, \\
F''_{zx} = R\dfrac{\partial \mu}{\partial x} + \mu\dfrac{\partial R}{\partial x} = F''_{xz} = P\dfrac{\partial \mu}{\partial z} + \mu\dfrac{\partial P}{\partial z}, & Q.
\end{array}$$

If these equations be multiplied by R, P, Q and added and if the result be simplified, the condition

$$P\left(\frac{\partial Q}{\partial z} - \frac{\partial R}{\partial y}\right) + Q\left(\frac{\partial R}{\partial x} - \frac{\partial P}{\partial z}\right) + R\left(\frac{\partial P}{\partial y} - \frac{\partial Q}{\partial x}\right) = 0 \qquad (5)$$

is found to be imposed on P, Q, R if there is to be an integrating factor. This is called the *condition of integrability*. For it may be shown conversely that if the condition (5) is satisfied, the equation may be integrated.

Suppose an attempt to integrate (1) be made as follows: First assume that one of the variables is constant (naturally, that one which will make the resulting equation simplest to integrate), say z. Then $Pdx + Qdy = 0$. Now integrate this simplified equation with an integrating factor or otherwise, and let $F(x, y, z) = \phi(z)$ be the integral, where the constant C is taken as a function ϕ of z. Next try to determine ϕ so that the integral $F(x, y, z) = \phi(z)$ will satisfy (1). To do this, differentiate;

$$F'_x dx + F'_y dy + F'_z dz = d\phi.$$

Compare this equation with (1). Then the equations*

$$F'_x = \lambda P, \qquad F'_y = \lambda Q, \qquad (F'_z - \lambda R)\,dz = d\phi$$

must hold. The third equation $(F'_z - \lambda R)\,dz = d\phi$ may be integrated provided the coefficient $S = F'_z - \lambda R$ of dz is a function of z and ϕ, that is, of z and F alone. This is so in case the condition (5) holds. It therefore appears that the integration of the equation (1) for which (5) holds reduces to the succession of two integrations of the type discussed in Chap. VIII.

As an example take $(2x^2 + 2xy + 2xz^2 + 1)\,dx + dy + 2zdz = 0$. The condition

$$(2x^2 + 2xy + 2xz^2 + 1)0 + 1(-4xz) + 2z(2x) = 0$$

of integrability is satisfied. The greatest simplification will be had by making x constant. Then $dy + 2zdz = 0$ and $y + z^2 = \phi(x)$. Compare

$$dy + 2zdz = d\phi \quad \text{and} \quad (2x^2 + 2xy + 2xz^2 + 1)\,dx + dy + 2zdz = 0.$$

Then $\lambda = 1, \quad -(2x^2 + 2xy + 2xz^2 + 1)\,dx = d\phi$;

or $-(2x^2 + 1 + 2x\phi)\,dx = d\phi$ or $d\phi + 2x\phi dx = -(2x^2 + 1)\,dx$.

This is the linear type with the integrating factor e^{x^2}. Then

$$e^{x^2}(d\phi + 2x\phi dx) = -e^{x^2}(2x^2 + 1)\,dx \quad \text{or} \quad e^{x^2}\phi = -\int e^{x^2}(2x^2 + 1)\,dx + C.$$

Hence $y + z^2 + e^{-x^2}\int e^{x^2}(2x^2 + 1)\,dx = Ce^{-x^2}$ or $e^{x^2}(y + z^2) + \int e^{x^2}(2x^2 + 1)\,dx = C$

is the solution. It may be noted that e^{x^2} is the integrating factor for the original equation:

$$e^{x^2}[(2x^2 + 2xy + 2xz^2 + 1)\,dx + dy + 2zdz] = d\left[e^{x^2}(y + z^2) + \int e^{x^2}(2x^2 + 1)\,dx\right].$$

To complete the proof that the equation (1) is integrable if (5) is satisfied, it is necessary to show that when the condition is satisfied the coefficient $S = F'_z - \lambda R$ is a function of z and F alone. Let it be regarded as a function of x, F, z instead of x, y, z. It is necessary to prove that the derivative of S by x when F and z are constant is zero. By the formulas for change of variable

$$\left(\frac{\partial S}{\partial x}\right)_{y,z} = \left(\frac{\partial S}{\partial x}\right)_{F,z} + \left(\frac{\partial S}{\partial F}\right)\frac{\partial F}{\partial x}, \qquad \left(\frac{\partial S}{\partial y}\right)_{x,z} = \left(\frac{\partial S}{\partial F}\right)_{x,z}\frac{\partial F}{\partial y}.$$

* Here the factor λ is not an integrating factor of (1), but only of the reduced equation $Pdx + Qdy = 0$.

But $F'_x = \lambda P$ and $F'_y = \lambda Q$, and hence $Q\left(\frac{\partial S}{\partial x}\right)_{y,z} - P\left(\frac{\partial S}{\partial y}\right)_{x,z} = Q\left(\frac{\partial S}{\partial x}\right)_{F,z}$.

Now $$\left(\frac{\partial S}{\partial x}\right)_{y,z} = \frac{\partial}{\partial x}\left(\frac{\partial F}{\partial z} - \lambda R\right) = \frac{\partial^2 F}{\partial z \partial x} - \frac{\partial \lambda R}{\partial x} = \frac{\partial \lambda P}{\partial z} - \frac{\partial \lambda R}{\partial x}.$$

Hence $$\left(\frac{\partial S}{\partial x}\right)_{y,z} = \lambda\left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}\right) + P\frac{\partial \lambda}{\partial z} - R\frac{\partial \lambda}{\partial x},$$

and $$\left(\frac{\partial S}{\partial y}\right)_{x,z} = \lambda\left(\frac{\partial Q}{\partial z} - \frac{\partial R}{\partial y}\right) + Q\frac{\partial \lambda}{\partial z} - R\frac{\partial \lambda}{\partial y}.$$

Then $$Q\left(\frac{\partial S}{\partial x}\right)_{y,z} - P\left(\frac{\partial S}{\partial y}\right)_{x,z} = \lambda\left[Q\left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}\right) + P\left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}\right)\right] - R\left[Q\frac{\partial \lambda}{\partial x} - P\frac{\partial \lambda}{\partial y}\right]$$

and $$Q\left(\frac{\partial S}{\partial x}\right)_{F,z} = \lambda\left[Q\left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}\right) + P\left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}\right) + R\left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)\right] - R\left[\frac{\partial \lambda Q}{\partial x} - \frac{\partial \lambda P}{\partial y}\right],$$

where a term has been added in the first bracket and subtracted in the second. Now as λ is an integrating factor for $Pdx + Qdy$, it follows that $(\lambda Q)'_x = (\lambda P)'_y$; and only the first bracket remains. By the condition of integrability this, too, vanishes and hence S as a function of x, F, z does not contain x but is a function of F and z alone, as was to be proved.

110. It has been seen that if the equation (1) is integrable, there is an integrating factor and the condition (5) is satisfied; also that conversely if the condition is satisfied the equation may be integrated. Geometrically this means that the infinity of planar elements defined by the equation can be grouped upon a family of surfaces $F(x, y, z) = C$ to which they are tangent. If the condition of integrability is not satisfied, the planar elements cannot be thus grouped into surfaces. Nevertheless if a surface $G(x, y, z) = 0$ be given, the planar element of (1) which passes through any point (x_0, y_0, z_0) of the surface will cut the surface $G = 0$ in a certain lineal element of the surface. Thus upon the surface $G(x, y, z) = 0$ there will be an infinity of lineal elements, one through each point, which satisfy the given equation (1). And these elements may be grouped into curves lying upon the surface. If the equation (1) is integrable, these curves will of course be the intersections of the given surface $G = 0$ with the surfaces $F = C$ defined by the integral of (1).

The method of obtaining the curves upon $G(x, y, z) = 0$ which are the integrals of (1), in case (5) does not possess an integral of the form $F(x, y, z) = C$, is as follows. Consider the two equations

$$Pdx + Qdy + Rdz = 0, \qquad G'_x dx + G'_y dy + G'_z dz = 0,$$

of which the first is the given differential equation and the second is the differential equation of the given surface. From these equations one of the differentials, say dz, may be eliminated, and the corresponding variable z may also be eliminated by substituting its value obtained by solving $G(x, y, z) = 0$. Thus there is obtained a differential equation $Mdx + Ndy = 0$ connecting the other two variables x and y. The integral of this, $F(x, y) = C$, consists of a family of cylinders which cut the given surface $G = 0$ in the curves which satisfy (1).

Consider the equation $ydx + xdy - (x + y + z)dz = 0$. This does not satisfy the condition (5) and hence is not completely integrable; but a set of integral curves may be found on any assigned surface. If the surface be the plane $z = x + y$, then

$$ydx + xdy - (x + y + z)dz = 0 \quad \text{and} \quad dz = dx + dy$$

give $\quad (x + z)dx + (y + z)dy = 0 \quad$ or $\quad (2x + y)dx + (2y + x)dy = 0$

by eliminating dz and z. The resulting equation is exact. Hence

$$x^2 + xy + y^2 = C \quad \text{and} \quad z = x + y$$

give the curves which satisfy the equation and lie in the plane.

If the equation (1) were integrable, the integral curves may be used to obtain the integral surfaces and thus to accomplish the complete integration of the equation by *Mayer's method*. For suppose that $F(x, y, z) = C$ were the integral surfaces and that $F(x, y, z) = F(0, 0, z_0)$ were that particular surface cutting the z-axis at z_0. The family of planes $y = \lambda x$ through the z-axis would cut the surface in a series of curves which would be integral curves, and the surface could be regarded as generated by these curves as the plane turned about the axis. To reverse these considerations let $y = \lambda x$ and $dy = \lambda dx$; by these relations eliminate dy and y from (1) and thus obtain the differential equation $Mdx + Ndz = 0$ of the intersections of the planes with the solutions of (1). Integrate the equation as $f(x, z, \lambda) = C$ and determine the constant so that $f(x, z, \lambda) = f(0, z_0, \lambda)$. For any value of λ this gives the intersection of $F(x, y, z) = F(0, 0, z_0)$ with $y = \lambda x$. Now if λ be eliminated by the relation $\lambda = y/x$, the result will be the surface

$$f\left(x, z, \frac{y}{x}\right) = f\left(0, z_0, \frac{y}{x}\right), \quad \text{equivalent to} \quad F(x, y, z) = F(0, 0, z_0),$$

which is the integral of (1) and passes through $(0, 0, z_0)$. As z_0 is arbitrary, the solution contains an arbitrary constant and is the general solution.

It is clear that instead of using planes through the z-axis, planes through either of the other axes might have been used, or indeed planes or cylinders through any line parallel to any of the axes. Such modifications are frequently necessary owing to the fact that the substitution $f(0, z_0, \lambda)$ introduces a division by 0 or a log 0 or some other impossibility. For instance consider

$$y^2dx + zdy - ydz = 0, \qquad y = \lambda x, \qquad dy = \lambda dx, \qquad \lambda^2x^2dx + \lambda zdx - \lambda xdz = 0.$$

Then $$\lambda dx + \frac{zdx - xdz}{x^2} = 0, \quad \text{and} \quad \lambda x - \frac{z}{x} = f(x, z, \lambda).$$

But here $f(0, z_0, \lambda)$ is impossible and the solution is illusory. If the planes $(y - 1) = \lambda x$ passing through a line parallel to the z-axis and containing the point $(0, 1, 0)$ had been used, the result would be

$$dy = \lambda dx, \qquad (1 + \lambda x)^2dx + \lambda zdx - (1 + \lambda x)dz = 0,$$

or $$dx + \frac{\lambda z dx - (1 + \lambda x) dz}{(1 + \lambda x)^2} = 0, \quad \text{and} \quad x - \frac{z}{1 + \lambda x} = f(x, z, \lambda).$$

Hence $$x - \frac{z}{1 + \lambda x} = - z_0 \quad \text{or} \quad x - \frac{z}{y} = - z_0 = C,$$

is the solution. The same result could have been obtained with $x = \lambda z$ or $y = \lambda (x - a)$. In the latter case, however, care should be taken to use $f(x, z, \lambda) = f(a, z_0, \lambda)$.

EXERCISES

1. Test these equations for exactness; if exact, integrate; if not exact, find an integrating factor by inspection and integrate:

(α) $(y + z) dx + (z + x) dy + (x + y) dz = 0$, ($\beta$) $y^2 dx + z dy - y dz = 0$,
(γ) $x dx + y dy - \sqrt{a^2 - x^2 - y^2} dz = 0$, ($\delta$) $2 z (dx - dy) + (x - y) dz = 0$,
(ϵ) $(2 x + y^2 + 2 xz) dx + 2 xy dy + x^2 dz = 0$, ($\zeta$) $zy dx = zx dy + y^2 dx$,
(η) $x (y - 1) (z - 1) dx + y (z - 1) (x - 1) dy + z (x - 1) (y - 1) dz = 0$.

2. Apply the test of integrability and integrate these:

(α) $(x^2 - y^2 - z^2) dx + 2 xy dy + 2 xz dz = 0$,
(β) $(x + y^2 + z^2 + 1) dx + 2 y dy + 2 z dz = 0$,
(γ) $(y + a)^2 dx + z dy = (y + a) dz$,
(δ) $(1 - x^2 - 2 y^2 z) dz = 2 xz dx + 2 yz^2 dy$,
(ϵ) $x^2 dx^2 + y^2 dy^2 - z^2 dz^2 + 2 xy dx dy = 0$,
(ζ) $z (x dx + y dy + z dz)^2 = (z^2 - x^2 - y^2) (x dx + y dy + z dz) dz$.

3. If the equation is homogeneous, the substitution $x = uz$, $y = vz$, frequently shortens the work. Show that if the given equation satisfies the condition of integrability, the new equation will satisfy the corresponding condition in the new variables and may be rendered exact by an obvious integrating factor. Integrate:

(α) $(y^2 + yz) dx + (xz + z^2) dy + (y^2 - xy) dz = 0$,
(β) $(x^2 y - y^3 - y^2 z) dx + (xy^2 - x^2 z - x^3) dy + (xy^2 + x^2 y) dz = 0$,
(γ) $(y^2 + yz + z^2) dx + (x^2 + xz + z^2) dy + (x^2 + xy + y^2) dz = 0$.

4. Show that (5) does not hold; integrate subject to the relation imposed:

(α) $y dx + x dy - (x + y + z) dz = 0$, $x + y + z = k$ or $y = kx$,
(β) $c (x dy + y dy) + \sqrt{1 - a^2 x^2 - b^2 y^2} dz = 0$, $a^2 x^2 + b^2 y^2 + c^2 z^2 = 1$,
(γ) $dz = ay dx + b dy$, $y = kx$ or $x^2 + y^2 + z^2 = 1$ or $y = f(x)$.

5. Show that if an equation is integrable, it remains integrable after any change of variables from x, y, z to u, v, w.

6. Apply Mayer's method to sundry of Exs. 2 and 3.

7. Find the conditions of exactness for an equation in four variables and write the formula for the integration. Integrate with or without a factor:

(α) $(2 x + y^2 + 2 xz) dx + 2 xy dy + x^2 dz + du = 0$,
(β) $yzu dx + xzu dy + xyu dz + xyz du = 0$,
(γ) $(y + z + u) dx + (x + z + u) dy + (x + y + u) dz + (x + y + z) du = 0$,
(δ) $u (y + z) dx + u (y + z + 1) dy + u dz - (y + z) du = 0$.

8. If an equation in four variables is integrable, it must be so when any one of the variables is held constant. Hence the four conditions of integrability obtained by writing (5) for each set of three coefficients must hold. Show that the conditions are satisfied in the following cases. Find the integrals by a generalization of the method in the text by letting one variable be constant and integrating the three remaining terms and determining the constant of integration as a function of the fourth in such a way as to satisfy the equations.

$(\alpha)\ z(y+z)\,dx + z(u-x)\,dy + y(x-u)\,dz + y(y+z)\,du = 0,$

$(\beta)\ uyzdx + uzx\log x dy + uxy\log x dz - xdu = 0.$

9. Try to extend the method of Mayer to such as the above in Ex. 8.

10. If $G(x, y, z) = a$ and $H(x, y, z) = b$ are two families of surfaces defining a family of curves as their intersections, show that the equation

$$(G'_y H'_z - G'_z H'_y)\,dx + (G'_z H'_x - G'_x H'_z)\,dy + (G'_x H'_y - G'_y H'_x)\,dz = 0$$

is the equation of the planar elements perpendicular to the curves at every point of the curves. Find the conditions on G and H that there shall be a family of surfaces which cut all these curves orthogonally. Determine whether the curves below have orthogonal trajectories (surfaces); and if they have, find the surfaces:

$(\alpha)\ y = x + a,\ z = x + b,$ $(\beta)\ y = ax + 1,\ z = bx,$

$(\gamma)\ x^2 + y^2 = a^2,\ z = b,$ $(\delta)\ xy = a,\ xz = b,$

$(\epsilon)\ x^2 + y^2 + z^2 = a^2,\ xy = b,$ $(\zeta)\ x^2 + 2y^2 + 3z^2 = a,\ xy + z = b,$

$(\eta)\ \log xy = az,\ x + y + z = b,$ $(\theta)\ y = 2ax + a^2,\ z = 2bx + b^2.$

11. Extend the work of proposition 3, § 94, and Ex. 11, p. 234, to find the normal derivative of the solution of equation (1) and to show that the singular solution may be looked for among the factors of $\mu^{-1} = 0$.

12. If $\mathbf{F} = P\mathbf{i} + Q\mathbf{j} + R\mathbf{k}$ be formed, show that (1) becomes $\mathbf{F}\cdot d\mathbf{r} = 0$. Show that the condition of exactness is $\nabla\times\mathbf{F} = 0$ by expanding $\nabla\times\mathbf{F}$ as the formal vector product of the operator ∇ and the vector $\mathbf{F}$ (see § 78). Show further that the condition of integrability is $\mathbf{F}\cdot(\nabla\times\mathbf{F}) = 0$ by similar formal expansion.

13. In Ex. 10 consider ∇G and ∇H. Show these vectors are normal to the surfaces $G = a$, $H = b$, and hence infer that $(\nabla G)\times(\nabla H)$ is the direction of the intersection. Finally explain why $d\mathbf{r}\cdot(\nabla G\times\nabla H) = 0$ is the differential equation of the orthogonal family if there be such a family. Show that this vector form of the family reduces to the form above given.

111. Systems of simultaneous equations.

The two equations

$$\frac{dy}{dx} = f(x, y, z), \qquad \frac{dz}{dx} = g(x, y, z) \tag{6}$$

in the two dependent variables y and z and the independent variable x constitute a set of simultaneous equations of the first order. It is more customary to write these equations in the form

$$\frac{dx}{X(x, y, z)} = \frac{dy}{Y(x, y, z)} = \frac{dz}{Z(x, y, z)}, \tag{7}$$

which is symmetric in the differentials and where $X : Y : Z = 1 : f : g$. At any assigned point x_0, y_0, z_0 of space the ratios $dx : dy : dz$ of the differentials are determined by substitution in (7). Hence the equations fix a definite direction at each point of space, that is, they determine a lineal element through each point. The problem of integration is to combine these lineal elements into a family of curves $F(x, y, z) = C_1$, $G(x, y, z) = C_2$, depending on two parameters C_1 and C_2, one curve passing through each point of space and having at that point the direction determined by the equations.

For the formal integration there are several allied methods of procedure. In the first place it may happen that two of

$$\frac{dx}{X} = \frac{dy}{Y}, \qquad \frac{dy}{Y} = \frac{dz}{Z}, \qquad \frac{dx}{X} = \frac{dz}{Z}.$$

are of such a form as to contain only the variables whose differentials enter. In this case these two may be integrated and the two solutions taken together give the family of curves. Or it may happen that one and only one of these equations can be integrated. Let it be the first and suppose that $F(x, y) = C_1$ is the integral. By means of this integral the variable x may be eliminated from the second of the equations or the variable y from the third. In the respective cases there arises an equation which may be integrated in the form $G(y, z, C_1) = C_2$ or $G(x, z, F) = C_2$, and this result taken with $F(x, y) = C_1$ will determine the family of curves.

Consider the example $\frac{xdx}{yz} = \frac{ydy}{xz} = \frac{dz}{y}$. Here the two equations

$$\frac{xdx}{y} = \frac{ydy}{x} \quad \text{and} \quad \frac{xdx}{z} = dz$$

are integrable with the results $x^3 - y^3 = C_1$, $x^2 - z^2 = C_2$, and these two integrals constitute the solution. The solution might, of course, appear in very different form; for there are an indefinite number of pairs of equations $F(x, y, z, C_1) = 0$, $G(x, y, z, C_2) = 0$ which will intersect in the curves of intersection of $x^3 - y^3 = C_1$, and $x^2 - z^2 = C_2$. In fact $(y^3 + C_1)^2 = (z^2 + C_2)^2$ is clearly a solution and could replace either of those found above.

Consider the example $\frac{dx}{x^2 - y^2 - z^2} = \frac{dy}{2xy} = \frac{dz}{2xz}$. Here

$$\frac{dy}{y} = \frac{dz}{z}, \quad \text{with the integral} \quad y = C_1 z,$$

is the only equation the integral of which can be obtained directly. If y be eliminated by means of this first integral, there results the equation

$$\frac{dx}{x^2 - (C_1^2 + 1) z^2} = \frac{dz}{2xz} \quad \text{or} \quad 2xzdx + [(C_1^2 + 1) z^2 - x^2] dz = 0.$$

This is homogeneous and may be integrated with a factor to give

$$x^2 + (C_1^2 + 1) z^2 = C_2 z \quad \text{or} \quad x^2 + y^2 + z^2 = C_2 z.$$

Hence $$y = C_1 z, \qquad x^2 + y^2 + z^2 = C_2 z$$

is the solution, and represents a certain family of circles.

Another method of attack is to use composition and division.

$$\frac{dx}{X}=\frac{dy}{Y}=\frac{dz}{Z}=\frac{\lambda dx+\mu dy+\nu dz}{\lambda X+\mu Y+\nu Z}. \tag{8}$$

Here λ, μ, ν may be chosen as any functions of (x, y, z). It may be possible so to choose them that the last expression, taken with one of the first three, gives an equation which may be integrated. With this first integral a second may be obtained as before. Or it may be that two different choices of λ, μ, ν can be made so as to give the two desired integrals. Or it may be possible so to select two sets of multipliers that the equation obtained by setting the two expressions equal may be solved for a first integral. Or it may be possible to choose λ, μ, ν so that the denominator $\lambda X+\mu Y+\nu Z=0$, and so that the numerator (which must vanish if the denominator does) shall give an equation

$$\lambda dx+\mu dy+\nu dz=0 \tag{9}$$

which satisfies the condition (5) of integrability and may be integrated by the methods of § 109.

Consider the equations $\frac{dx}{x^2+y^2+yz}=\frac{dy}{x^2+y^2-xz}=\frac{dz}{(x+y)z}$. Here take λ, μ, ν as $1, -1, -1$; then $\lambda X+\mu Y+\nu Z=0$ and $dx-dy-dz=0$ is integrable as $x-y-z=C_1$. This may be used to obtain another integral. But another choice of λ, μ, ν as x, y, 0, combined with the last expression, gives

$$\frac{xdx+ydz}{(x^2+y^2)(x+y)}=\frac{dz}{(x+y)z} \quad \text{or} \quad \log(x^2+y^2)=\log z^2+C_2.$$

Hence $$x-y-z=C_1 \quad \text{and} \quad x^2+y^2=C_2z^2$$

will serve as solutions. This is shorter than the method of elimination.

It will be noted that these equations just solved are homogeneous. The substitution $x=uz$, $y=vz$ might be tried. Then

$$\frac{udz+zdu}{u^2+v^2+v}=\frac{vdz+zdv}{u^2+v^2-u}=\frac{dz}{u+v}=\frac{zdu}{v^2-uv+v}=\frac{zdv}{u^2-uv-u},$$

or $$\frac{du}{v^2-uv+v}=\frac{dv}{u^2-uv-u}=\frac{dz}{z}.$$

Now the first equations do not contain z and may be solved. This always happens in the homogeneous case and may be employed if no shorter method suggests itself.

It need hardly be mentioned that all these methods apply equally to the case where there are more than three equations. The geometric picture, however, fails, although the geometric language may be continued if one wishes to deal with higher dimensions than three. In some cases the introduction of a fourth variable, as

$$\frac{dx}{X}=\frac{dy}{Y}=\frac{dz}{Z}=\frac{dt}{1} \quad \text{or} \quad =\frac{dt}{t}, \tag{10}$$

is useful in solving a set of equations which originally contained only three variables. This is particularly true when X, Y, Z are linear with constant coefficients, in which case the methods of § 98 may be applied with t as independent variable.

112. Simultaneous differential equations of higher order, as

$$\frac{d^2x}{dt^2} = X\left(x,\, y,\, \frac{dx}{dt}\, \frac{dy}{dt}\right), \qquad \frac{d^2y}{dt^2} = Y\left(x,\, y,\, \frac{dx}{dt}\, \frac{dy}{dt}\right),$$

$$\frac{d^2r}{dt^2} - \frac{1}{r}\left(\frac{d\phi}{dt}\right)^2 = R\left(r,\, \phi,\, \frac{dr}{dt},\, \frac{d\phi}{dt}\right), \qquad \frac{1}{r}\frac{d}{dt}\left(r^2\frac{d\phi}{dt}\right) = \Phi\left(r,\, \phi,\, \frac{dr}{dt},\, \frac{d\phi}{dt}\right),$$

especially those of the second order like these, are of constant occurrence in mechanics; for the acceleration requires second derivatives with respect to the time for its expression, and the forces are expressed in terms of the coördinates and velocities. The complete integration of such equations requires the expression of the dependent variables as functions of the independent variable, generally the time, with a number of constants of integration equal to the sum of the orders of the equations. Frequently even when the complete integrals cannot be found, it is possible to carry out some integrations and replace the given system of equations by fewer equations or equations of lower order containing some constants of integration.

No special or general rules will be laid down for the integration of systems of higher order. In each case some particular combinations of the equations may suggest themselves which will enable an integration to be performed.* In problems in mechanics the principles of energy, momentum, and moment of momentum frequently suggest combinations leading to integrations. Thus if

$$x'' = X, \qquad y'' = Y, \qquad z'' = Z,$$

where accents denote differentiation with respect to the time, be multiplied by dx, dy, dz and added, the result

$$x''dx + y''dy + z''dz = Xdx + Ydy + Zdz \tag{11}$$

contains an exact differential on the left; then if the expression on the right is an exact differential, the integration

$$\tfrac{1}{2}(x'^2 + y'^2 + z'^2) = \int Xdx + Ydy + Zdz + C \tag{11'}$$

* It is possible to differentiate the given equations repeatedly and eliminate all the dependent variables except one. The resulting differential equation, say in x and t, may then be treated by the methods of previous chapters; but this is rarely successful except when the equation is linear.

can be performed. This is *the principle of energy* in its simplest form. If two of the equations are multiplied by the chief variable of the other and subtracted, the result is

$$yx'' - xy'' = yX - xY \tag{12}$$

and the expression on the left is again an exact differential; if the right-hand side reduces to a constant or a function of t, then

$$yx' - xy' = \int f(t) + C \tag{12'}$$

is an integral of the equations. This is *the principle of moment of momentum.* If the equations can be multiplied by constants as

$$lx'' + my'' + nz'' = lX + mY + nZ, \tag{13}$$

so that the expression on the right reduces to a function of t, an integration may be performed. This is *the principle of momentum.* These three are the most commonly usable devices.

As an example: Let a particle move in a plane subject to forces attracting it toward the axes by an amount proportional to the mass and to the distance from the axes; discuss the motion. Here the equations of motion are merely

$$m\frac{d^2x}{dt^2} = -kmx, \qquad m\frac{d^2y}{dt^2} = -kmy \quad \text{or} \quad \frac{d^2x}{dt^2} = -kx, \qquad \frac{d^2y}{dt^2} = -ky.$$

Then $$x\frac{d^2x}{dt^2} + y\frac{d^2y}{dt^2} = -k(xdx + ydy) \quad \text{and} \quad \left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 = -(x^2 + y^2) + C.$$

Also $$y\frac{d^2x}{dt^2} - x\frac{d^2y}{dt^2} = 0 \quad \text{and} \quad y\frac{dx}{dt} - x\frac{dy}{dt} = C'.$$

In this case the two principles of energy and moment of momentum give two integrals and the equations are reduced to two of the first order. But as it happens, the original equations could be integrated directly as

$$\frac{d^2x}{dt^2}dx = -kxdx, \qquad \left(\frac{dx}{dt}\right)^2 = -kx^2 + C^2, \qquad \frac{dx}{\sqrt{C^2 - kx^2}} = dt$$

$$\frac{d^2y}{dt^2}dy = -kydy, \qquad \left(\frac{dy}{dt}\right)^2 = -ky^2 + K^2, \qquad \frac{dy}{\sqrt{K^2 - ky^2}} = dt.$$

The constants C^2 and K^2 of integration have been written as squares because they are necessarily positive. The complete integration gives

$$\sqrt{k}x = C\sin\left(\sqrt{k}t + C_1\right), \qquad \sqrt{k}y = K\sin\left(\sqrt{k}t + K_2\right).$$

As another example: A particle, attracted toward a point by a force equal to $r/m^2 + h^2/r^3$ per unit mass, where m is the mass and h is the double areal velocity and r is the distance from the point, is projected perpendicularly to the radius vector at the distance $\sqrt{mh}$; discuss the motion. In polar coördinates the equations of motion are

$$m\left[\frac{d^2r}{dt^2} - \frac{1}{r}\left(\frac{d\phi}{dt}\right)^2\right] = R = -\frac{mr}{m^2} - \frac{mh^2}{r^3}, \qquad \frac{m}{r}\frac{d}{dt}\left(r^2\frac{d\phi}{dt}\right) = \Phi = 0.$$

The second integrates directly as $r^2 d\phi/dt = h$ where the constant of integration h is twice the areal velocity. Now substitute in the first to eliminate ϕ.

$$\frac{d^2r}{dt^2} - \frac{h^2}{r^3} = -\frac{r}{m^2} - \frac{h^2}{r^3} \quad \text{or} \quad \frac{d^2r}{dt^2} = -\frac{r}{m^2} \quad \text{or} \quad \left(\frac{dr}{dt}\right)^2 = -\frac{r^2}{m^2} + C.$$

Now as the particle is projected perpendicularly to the radius, $dr/dt = 0$ at the start when $r = \sqrt{mh}$. Hence the constant C is h/m. Then

$$\frac{dr}{\sqrt{\frac{h}{m} - \frac{r^2}{m^2}}} = dt \quad \text{and} \quad \frac{r^2 d\phi}{h} = dt \quad \text{give} \quad \frac{\sqrt{mh}\,dr}{r^2\sqrt{1 - \frac{r^2}{hm}}} = d\phi.$$

Hence $$\sqrt{mh}\sqrt{\frac{1}{r^2} - \frac{1}{h}} = \phi + C \quad \text{or} \quad \frac{1}{r^2} - \frac{1}{hm} = \frac{(\phi + C)^2}{mh}.$$

Now if it be assumed that $\phi = 0$ at the start when $r = \sqrt{mh}$, we find $C = 0$.

Hence $$r^2 = \frac{mh}{1 + \phi^2} \quad \text{is the orbit.}$$

To find the relation between ϕ and the time,

$$r^2 d\phi = h dt \quad \text{or} \quad \frac{m d\phi}{1 + \phi^2} = dt \quad \text{or} \quad t = m \tan^{-1}\phi,$$

if the time be taken as $t = 0$ when $\phi = 0$. Thus the orbit is found, the expression of ϕ as a function of the time is found, and the expression of r as a function of the time is obtainable. The problem is completely solved. It will be noted that the constants of integration have been determined after each integration by the initial conditions. This simplifies the subsequent integrations which might in fact be impossible in terms of elementary functions without this simplification.

EXERCISES

1. Integrate these equations:

(α) $\frac{dx}{yz} = \frac{dy}{xz} = \frac{dz}{xy}$, (β) $\frac{dx}{y^2} = \frac{dy}{x^2} = \frac{dz}{x^2y^2z^2}$,

(γ) $\frac{dx}{xz} = \frac{dy}{yz} = \frac{dz}{xy}$, (δ) $\frac{dx}{yz} = \frac{dy}{xz} = \frac{dz}{x + y}$,

(ϵ) $-\frac{dx}{y} = \frac{dy}{x} = \frac{dz}{1 + z^2}$, (ζ) $\frac{dx}{-1} = \frac{dy}{3y + 4z} = \frac{dz}{2y + 5z}$.

2. Integrate the equations: (α) $\frac{dx}{bz - cy} = \frac{dy}{cx - az} = \frac{dz}{ay - bx}$,

(β) $\frac{dx}{x^2 + y^2} = \frac{dy}{2xy} = \frac{dz}{xz + yz}$, (γ) $\frac{dx}{y + z} = \frac{dy}{x + z} = \frac{dz}{x + y}$,

(δ) $\frac{dx}{y^3x - 2x^4} = \frac{dy}{2y^4 - x^3y} = \frac{dz}{yz(x^3 - y^3)}$, (ϵ) $\frac{dx}{x(y - z)} = \frac{dy}{y(z - x)} = \frac{dz}{z(x - y)}$,

(ζ) $\frac{dx}{x(y^2 - z^2)} = \frac{dy}{y(z^2 - x^2)} = \frac{dz}{z(x^2 - y^2)}$, (η) $\frac{dx}{x(y^2 - z^2)} = \frac{-dy}{y(z^2 + x^2)} = \frac{dz}{z(x^2 + y^2)}$,

(θ) $\frac{dx}{y - z} = \frac{dy}{x + y} = \frac{dz}{x + z} = dt$, (ι) $\frac{dx}{y - z} = \frac{dy}{x + y + t} = \frac{dz}{x + z + t} = dt$.

3. Show that the differential equations of the orthogonal trajectories (curves of the family of surfaces $F(x, y, z) = C$ are $dx : dy : dz = F'_x : F'_y : F'_z$. Find the curves which cut the following families of surfaces orthogonally:

(α) $a^2x^2 + b^2y^2 + c^2z^2 = C$, (β) $xyz = C$, (γ) $y^2 = Cxz$,
(δ) $y = x\tan(z + C)$, (ϵ) $y = x\tan Cz$, (ζ) $z = Cxy$.

4. Show that the solution of $dx : dy : dz = X : Y : Z$, where X, Y, Z are linear expressions in x, y, z, can always be found provided a certain cubic equation can be solved.

5. Show that the solutions of the two equations

$$\frac{dx}{dt} + T(ax + by) = T_1, \qquad \frac{dy}{dt} + T(a'x + b'y) = T_2,$$

where T, T_1, T_2 are functions of t, may be obtained by adding the equation as

$$\frac{d}{dt}(x + ly) + \lambda T(x + ly) = T_1 + lT_2$$

after multiplying one by l, and by determining λ as a root of

$$\lambda^2 - (a + b')\lambda + ab' - a'b = 0.$$

6. Solve: (α) $t\dfrac{dx}{dt} + 2(x - y) = t, \quad t\dfrac{dy}{dt} + x + 5y = t^2,$

(β) $tdx = (t - 2x)\,dt, \quad tdy = (tx + ty + 2x - t)\,dt,$

(γ) $\dfrac{ldx}{mn(y - z)} = \dfrac{mdy}{nl(z - x)} = \dfrac{ndz}{lm(x - y)} = \dfrac{dt}{t}.$

7. A particle moves in vacuo in a vertical plane under the force of gravity alone. Integrate. Determine the constants if the particle starts from the origin with a velocity V and at an angle of α degrees with the horizontal and at the time $t = 0$.

8. Same problem as in Ex. 7 except that the particle moves in a medium which resists proportionately to the velocity of the particle.

9. A particle moves in a plane about a center of force which attracts proportionally to the distance from the center and to the mass of the particle.

10. Same as Ex. 9 but with a repulsive force instead of an attracting force.

11. A particle is projected parallel to a line toward which it is attracted with a force proportional to the distance from the line.

12. Same as Ex. 11 except that the force is inversely proportional to the square of the distance and only the path of the particle is wanted.

13. A particle is attracted toward a center by a force proportional to the square of the distance. Find the orbit.

14. A particle is placed at a point which repels with a constant force under which the particle moves away to a distance a where it strikes a peg and is deflected off at a right angle with undiminished velocity. Find the orbit of the subsequent motion.

15. Show that equations (7) may be written in the form $d\mathbf{r}\times\mathbf{F} = 0$. Find the condition on $\mathbf{F}$ or on X, Y, Z that the integral curves have orthogonal surfaces.

113. Introduction to partial differential equations. An equation which contains a dependent variable, two or more independent variables, and one or more partial derivatives of the dependent variable with respect to the independent variables is called a *partial differential equation.* The equation

$$P(x, y, z)\frac{\partial z}{\partial x} + Q(x, y, z)\frac{\partial z}{\partial y} = R(x, y, z), \quad p = \frac{\partial z}{\partial x}, \quad q = \frac{\partial z}{\partial y}, \tag{14}$$

is clearly a linear partial differential equation of the first order in one dependent and two independent variables. The discussion of this equation preliminary to its integration may be carried on by means of the concept of *planar elements,* and the discussion will immediately suggest the method of integration.

When any point (x_0, y_0, z_0) of space is given, the coefficients P, Q, R in the equation take on definite values and the derivatives p and q are connected by a linear relation. Now any planar element through (x_0, y_0, z_0) may be considered as specified by the two slopes p and q; for it is an infinitesimal portion of the plane $z - z_0 = p(x - x_0) + q(y - y_0)$ in the neighborhood of the point. This plane contains the line or lineal element whose direction is

$$dx : dy : dz = P : Q : R, \tag{15}$$

because the substitution of P, Q, R for $dx = x - x_0$, $dy = y - y_0$, $dz = z - z_0$ in the plane gives the original equation $Pp + Qq = R$. Hence it appears that the planar elements defined by (14), of which there are an infinity through each point of space, are so related that all which pass through a given point of space pass through a certain line through that point, namely the line (15).

Now the problem of integrating the equation (14) is that of grouping the planar elements which satisfy it into surfaces. As at each point they are already grouped in a certain way by the lineal elements through which they pass, it is first advisable to group these lineal elements into curves by integrating the simultaneous equations (15). The integrals of these equations are the curves defined by two families of surfaces $F(x, y, z) = C_1$ and $G(x, y, z) = C_2$. These curves are called the *characteristic curves* or merely the *characteristics* of the equation (14). Through each lineal element of these curves there pass an infinity of the planar elements which satisfy (14). It is therefore clear that if these curves be in any wise grouped into surfaces, the planar elements of the surfaces must satisfy (14); for through each point of the surfaces will pass one of the curves, and the planar element of the surface at that point must therefore pass through the lineal element of the curve and hence satisfy (14).

To group the curves $F(x, y, z) = C_1$, $G(x, y, z) = C_2$ which depend on two parameters C_1, C_2 into a surface, it is merely necessary to introduce some functional relation $C_2 = f(C_1)$ between the parameters so that when one of them, as C_1, is given, the other is determined, and thus a particular curve of the family is fixed by one parameter alone and will sweep out a surface as the parameter varies. Hence *to integrate* (14), *first integrate* (15) *and then write*

$$G(x, y, z) = \Phi[F(x, y, z)] \quad \text{or} \quad \Phi(F, G) = 0, \tag{16}$$

where Φ denotes any arbitrary function. This will be the integral of (14) and will contain an arbitrary function Φ.

As an example, integrate $(y - z)p + (z - x)q = x - y$. Here the equations

$$\frac{dx}{y - z} = \frac{dy}{z - x} = \frac{dz}{x - y} \quad \text{give} \quad x^2 + y^2 + z^2 = C_1, \quad x + y + z = C_2$$

as the two integrals. Hence the solution of the given equation is

$$x + y + z = \Phi(x^2 + y^2 + z^2) \quad \text{or} \quad \Phi(x^2 + y^2 + z^2, x + y + z) = 0,$$

where Φ denotes an arbitrary function. The arbitrary function allows a solution to be determined which shall pass through any desired curve; for if the curve be $f(x, y, z) = 0$, $g(x, y, z) = 0$, the elimination of x, y, z from the four simultaneous equations

$$F(x, y, z) = C_1, \qquad G(x, y, z) = C_2, \qquad f(x, y, z) = 0, \qquad g(x, y, z) = 0$$

will express the condition that the four surfaces meet in a point, that is, that the curve given by the first two will cut that given by the second two; and this elimination will determine a relation between the two parameters C_1 and C_2 which will be precisely the relation to express the fact that the integral curves cut the given curve and that consequently the surface of integral curves passes through the given curve. Thus in the particular case here considered, suppose the solution were to pass through the curve $y = x^2$, $z = x$; then

$$x^2 + y^2 + z^2 = C_1, \qquad x + y + z = C_2, \qquad y = x^2, \qquad z = x$$

give
$$2x^2 + x^4 = C_1, \qquad x^2 + 2x = C_2,$$

whence
$$(C_2^2 + 2C_2 - C_1)^2 + 8C_2^2 - 24C_1 - 16C_1C_2 = 0.$$

The substitution of $C_1 = x^2 + y^2 + z^2$ and $C_2 = x + y + z$ in this equation will give the solution of $(y - z)p + (z - x)q = x - y$ which passes through the parabola $y = x^2$, $z = x$.

114. It will be recalled that the integral of an ordinary differential equation $f(x, y, y', \cdots, y^{(n)}) = 0$ of the nth order contains n constants, and that conversely if a system of curves in the plane, say $F(x, y, C_1, \cdots, C_n) = 0$, contains n constants, the constants may be eliminated from the equation and its first n derivatives with respect to x. It has now been seen that the integral of a certain partial differential equation contains an arbitrary function, and it might be inferred that the elimination of an arbitrary function would give rise to a partial differential equation of the first order. To show this, suppose $F(x, y, z) = \Phi[G(x, y, z)]$. Then

$$F'_x + F'_z p = \Phi' \cdot (G'_x + G'_z p), \qquad F'_y + F'_z q = \Phi' \cdot (G'_y + G'_z q)$$

follow from partial differentiation with respect to x and y; and

$$(F'_z G'_y - F'_y G'_z)p + (F'_x G'_z - F'_z G'_x) q = F'_y G'_x - F'_x G'_y$$

is a partial differential equation arising from the elimination of Φ'. More generally, the elimination of n arbitrary functions will give rise to an equation of the nth order; conversely it may be believed that the integration of such an equation would introduce n arbitrary functions in the general solution.

As an example, eliminate from $z = \Phi(xy) + \Psi(x + y)$ the two arbitrary functions Φ and Ψ. The first differentiation gives

$$p = \Phi' \cdot y + \Psi', \qquad q = \Phi' \cdot x + \Psi', \qquad p - q = (y - x)\Phi'.$$

Now differentiate again and let $r = \dfrac{\partial^2 z}{\partial x^2}$, $s = \dfrac{\partial^2 z}{\partial x \partial y}$, $t = \dfrac{\partial^2 z}{\partial y^2}$. Then

$$r - s = -\Phi' + (y - x)\Phi'' \cdot y, \qquad s - t = \Phi' + (y - x)\Phi'' \cdot x.$$

These two equations with $p - q = (y - x)\Phi'$ make three from which

$$xr - (x + y)s + yt = \frac{x + y}{x - y}(p - q) \quad \text{or} \quad x\frac{\partial^2 z}{\partial x^2} - (x + y)\frac{\partial^2 z}{\partial x \partial y} + y\frac{\partial^2 z}{\partial y^2} = \frac{x + y}{x - y}\left(\frac{\partial z}{\partial x} - \frac{\partial z}{\partial y}\right)$$

may be obtained as a partial differential equation of the second order free from Φ and Ψ. The general integral of this equation would be $z = \Phi(xy) + \Psi(x + y)$.

A partial differential equation may represent a certain definite type of surface. For instance by definition a conoidal surface is a surface generated by a line which moves parallel to a given plane, the director plane, and cuts a given line, the directrix. If the director plane be taken as $z = 0$ and the directrix be the z-axis, the equations of any line of the surface are

$$z = C_1, \qquad y = C_2 x, \qquad \text{with} \quad C_1 = \Phi(C_2)$$

as the relation which picks out a definite family of the lines to form a particular conoidal surface. Hence $z = \Phi(y/x)$ may be regarded as the general equation of a conoidal surface of which $z = 0$ is the director plane and the z-axis the directrix. The elimination of Φ gives $px + qy = 0$ as the differential equation of any such conoidal surface.

Partial differentiation may be used not only to eliminate arbitrary functions, but to eliminate constants. For if an equation $f(x, y, z, C_1, C_2) = 0$ contained two constants, the equation and its first derivatives with respect to x and y would yield three equations from which the constants could be eliminated, leaving a partial differential equation $F(x, y, z, p, q) = 0$ of the first order. If there had been five constants, the equation with its two first derivatives and its three second derivatives with respect to x and y would give a set of six equations from which the constants could be eliminated, leaving a differential equation of the second order. And so on. As the differential equation is obtained by eliminating the constants, the original equation will be a solution of the resulting differential equation.

For example, eliminate from $z = Ax^2 + 2Bxy + Cy^2 + Dx + Ey$ the five constants. The two first and three second derivatives are

$$p = 2Ax + 2By + D, \quad q = 2Bx + 2Cy + E, \quad r = 2A, \quad s = 2B, \quad t = 2C.$$

Hence
$$z = -\tfrac{1}{2}rx^2 - \tfrac{1}{2}ty^2 - sxy + px + qy$$

is the differential equation of the family of surfaces. The family of surfaces do not constitute the general solution of the equation, for that would contain two arbitrary functions, but they give what is called a *complete solution.* If there had been only three or four constants, the elimination would have led to a differential equation of the second order which need have contained only one or two of the second derivatives instead of all three; it would also have been possible to find three or two simultaneous partial differential equations by differentiating in different ways.

115. If $f(x, y, z, C_1, C_2) = 0$ and $F(x, y, z, p, q) = 0$ (17)

are two equations of which the second is obtained by the elimination of the two constants from the first, the first is said to be the *complete solution* of the second. That is, any equation which contains two distinct arbitrary constants and which satisfies a partial differential equation of the first order is said to be a complete solution of the differential equation. A complete solution has an interesting geometric interpretation. The differential equation $F = 0$ defines a series of planar elements through each point of space. So does $f(x, y, z, C_1, C_2) = 0$. For the tangent plane is given by

$$\left.\frac{\partial f}{\partial x}\right|_0 (x - x_0) + \left.\frac{\partial f}{\partial y}\right|_0 (y - y_0) + \left.\frac{\partial f}{\partial z}\right|_0 (z - z_0) = 0$$

with
$$f(x_0, y_0, z_0, C_1, C_2) = 0$$

as the condition that C_1 and C_2 shall be so related that the surface passes through (x_0, y_0, z_0). As there is only this one relation between the two arbitrary constants, there is a whole series of planar elements through the point. As $f(x, y, z, C_1, C_2) = 0$ satisfies the differential equation, the planar elements defined by it are those defined by the differential equation. Thus a complete solution establishes an arrangement of the planar elements defined by the differential equation upon a family of surfaces dependent upon two arbitrary constants of integration.

From the idea of a solution of a partial differential equation of the first order as a surface pieced together from planar elements which satisfy the equation, it appears that the envelope (p. 140) of any family of solutions will itself be a solution; for each point of the envelope is a point of tangency with some one of the solutions of the family, and the planar element of the envelope at that point is identical with the planar element of the solution and hence satisfies the differential equation. *This observation allows the general solution to be determined from any complete solution.* For if in $f(x, y, z, C_1, C_2) = 0$ any relation $C_2 = \Phi(C_1)$ is introduced between the two arbitrary constants, there arises a family depending on one parameter, and the envelope of the family is found by eliminating C_1 from the three equations

$$C_2 = \Phi(C_1), \qquad \frac{\partial f}{\partial C_1} + \frac{d\Phi}{dC_1}\frac{\partial f}{\partial C_2} = 0, \qquad f = 0. \tag{18}$$

As the relation $C_2 = \Phi(C_1)$ contains an arbitrary function Φ, the result of the elimination may be considered as containing an arbitrary function even though it is generally impossible to carry out the elimination except in the case where Φ has been assigned and is therefore no longer arbitrary.

A family of surfaces $f(x, y, z, C_1, C_2) = 0$ depending on two parameters may also have an envelope (p. 139). This is found by eliminating C_1 and C_2 from the three equations

$$f(x, y, z, C_1, C_2) = 0, \qquad \frac{\partial f}{\partial C_1} = 0, \qquad \frac{\partial f}{\partial C_2} = 0.$$

This surface is tangent to all the surfaces in the complete solution. This envelope is called the ***singular solution*** of the partial differential equation. As in the case of ordinary differential equations (§ 101), the singular solution may be obtained directly from the equation;* it is merely necessary to eliminate p and q from the three equations

$$F(x, y, z, p, q) = 0, \qquad \frac{\partial F}{\partial p} = 0, \qquad \frac{\partial F}{\partial q} = 0.$$

The last two equations express the fact that $F(p, q) = 0$ regarded as a function of p and q should have a double point (§ 57). A reference to § 67 will bring out another point, namely, that not only are all the surfaces represented by the complete solution tangent to the singular solution, but so is any surface which is represented by the general solution.

* It is hardly necessary to point out the fact that, as in the case of ordinary equations, extraneous factors may arise in the elimination, whether of C_1, C_2 or of p, q.

EXERCISES

1. Integrate these linear equations:

(α) $xzp + yzq = xy$, (β) $a(p+q) = z$, (γ) $x^2p + y^2q = z^2$,
(δ) $-yp + xq + 1 + z^2 = 0$, (ϵ) $yp - xq = x^2 - y^2$, (ζ) $(x+z)p = y$,
(η) $x^2p - xyq + y^2 = 0$, (θ) $(a-x)p + (b-y)q = c - z$,
(ι) $p\tan x + q\tan y = \tan z$, ($\kappa$) $(y^2 + z^2 - x^2)p - 2xyq + 2xz = 0$.

2. Determine the integrals of the preceding equations to pass through the curves:

for (α) $x^2 + y^2 = 1,\ z = 0$, for (β) $y = 0,\ x = z$,
for (γ) $y = 2x,\ z = 1$, for (ϵ) $x = z,\ y = z$.

3. Show analytically that if $F(x, y, z) = C_1$ is a solution of (15), it is a solution of (14). State precisely what is meant by a solution of a partial differential equation, that is, by the statement that $F(x, y, z) = C_1$ satisfies the equation. Show that the equations

$$P\frac{\partial z}{\partial x} + Q\frac{\partial z}{\partial y} = R \quad \text{and} \quad P\frac{\partial F}{\partial x} + Q\frac{\partial F}{\partial y} + R\frac{\partial F}{\partial z} = 0$$

are equivalent and state what this means. Show that if $F = C_1$ and $G = C_2$ are two solutions, then $F = \Phi(G)$ is a solution, and show conversely that a functional relation must exist between any two solutions (see § 62).

4. Generalize the work in the text along the analytic lines of Ex. 3 to establish the rules for integrating a linear equation in one dependent and four or n independent variables. In particular show that the integral of

$$P_1\frac{\partial z}{\partial x_1} + \cdots + P_n\frac{\partial z}{\partial x_n} = P_{n+1} \quad \text{depends on} \quad \frac{dx_1}{P_1} = \cdots = \frac{dx_n}{P_n} = \frac{dz}{P_{n+1}},$$

and that if $F_1 = C_1, \cdots, F_n = C_n$ are n integrals of the simultaneous system, the integral of the partial differential equation is $\Phi(F_1, \cdots, F_n) = 0$.

5. Integrate: (α) $x\dfrac{\partial u}{\partial x} + y\dfrac{\partial u}{\partial y} + z\dfrac{\partial u}{\partial z} = xyz$,

(β) $(y + z + u)\dfrac{\partial u}{\partial x} + (z + u + x)\dfrac{\partial u}{\partial y} + (u + x + y)\dfrac{\partial u}{\partial z} = x + y + z$.

6. Interpret the general equation of the first order $F(x, y, z, p, q) = 0$ as determining at each point (x_0, y_0, z_0) of space a series of planar elements tangent to a certain cone, namely, the cone found by eliminating p and q from the three simultaneous equations

$$F(x_0, y_0, z_0, p, q) = 0, \qquad (x - x_0)p + (y - y_0)q = z - z_0,$$
$$(x - x_0)\frac{\partial F}{\partial q} - (y - y_0)\frac{\partial F}{\partial p} = 0.$$

7. Eliminate the arbitrary functions:

(α) $x + y + z = \Phi(x^2 + y^2 + z^2)$, ($\beta$) $\Phi(x^2 + y^2, z - xy) = 0$,
(γ) $z = \Phi(x + y) + \Psi(x - y)$, ($\delta$) $z = e^{ny}\Phi(x - y)$,
(ϵ) $z = y^2 + 2\Phi(x^{-1} + \log y)$, ($\zeta$) $\Phi\left(\dfrac{x}{y}, \dfrac{y}{z}, \dfrac{z}{x}\right) = 0$.

8. Find the differential equations of these types of surfaces:

(α) cylinders with generators parallel to the line $x = az$, $y = bz$,

(β) conical surfaces with vertex at (a, b, c),

(γ) surfaces of revolution about the line $x : y : z = a : b : c$.

9. Eliminate the constants from these equations:

(α) $z = (x + a)(y + b)$, (β) $a(x^2 + y^2) + bz^2 = 1$,

(γ) $(x - a)^2 + (y - b)^2 + (z - c)^2 = 1$, ($\delta$) $(x - a)^2 + (y - b)^2 + (z - c)^2 = d^2$,

(ϵ) $Ax^2 + Bxy + Cy^2 + Dxz + Eyz = z^2$.

10. Show geometrically and analytically that $F(x, y, z) + aG(x, y, z) = b$ is a complete solution of the linear equation.

11. How many constants occur in the complete solution of the equation of the third, fourth, or nth order?

12. Discuss the complete, general, and singular solutions of an equation of the first order $F(x, y, z, u, u'_x, u'_y, u'_z) = 0$ with three independent variables.

13. Show that the planes $z = ax + by + C$, where a and b are connected by the relation $F(a, b) = 0$, are complete solutions of the equation $F(p, q) = 0$. Integrate:

(α) $pq = 1$, (β) $q = p^2 + 1$, (γ) $p^2 + q^2 = m^2$,

(δ) $pq = k$, (ϵ) $k \log q + p = 0$, (ζ) $3p^2 - 2q^2 = 4pq$,

and determine also the singular solutions.

14. Note that a simple change of variable will often reduce an equation to the type of Ex. 13. Thus the equations

$$F\left(\frac{p}{z}, \frac{q}{z}\right) = 0, \quad F(xp, q) = 0, \quad F\left(\frac{xp}{z}, \frac{yq}{z}\right) = 0,$$

with $\quad z = e^{z'}, \quad x = e^{x'}, \quad z = e^{z'}, x = e^{x'}, y = e^{y'},$

take a simpler form. Integrate and determine the singular solutions:

(α) $q = z + px$, (β) $x^2p^2 + y^2q^2 = z^2$, (γ) $z = pq$,

(δ) $q = 2yp^2$, (ϵ) $(p - y)^2 + (q - x)^2 = 1$, ($\zeta$) $z = p^m q^m$.

15. What is the obvious complete solution of the extended Clairaut equation $z = xp + yq + f(p, q)$? Discuss the singular solution. Integrate the equations:

(α) $z = xp + yq + \sqrt{p^2 + q^2 + 1}$, ($\beta$) $z = xp + yq + (p + q)^2$,

(γ) $z = xp + yq + pq$, (δ) $z = xp + yq - 2\sqrt{pq}$.

116. Types of partial differential equations. In addition to the linear equation and the types of Exs. 13–15 above, there are several types which should be mentioned. Of these the first is *the general equation of the first order*. If $F(x, y, z, p, q) = 0$ is the given equation and if a second equation $\Phi(x, y, z, p, q, a) = 0$, which holds simultaneously with the first and contains an arbitrary constant can be found, the two equations may be solved together for the values of p and q, and the results may be substituted in the relation $dz = pdx + qdy$ to give a total differential equation of which the integral will contain the constant a and a second constant of integration b. This integral will then be a complete integral of the given equation; the general integral may then be obtained by (18) of § 115. This is known as *Charpit's method.*

To find a relation $\Phi = 0$ differentiate the two equations

$$F(x, y, z, p, q) = 0, \qquad \Phi(x, y, z, p, q, a) = 0 \tag{19}$$

with respect to x and y and use the relation that dz be exact.

$$\begin{array}{r|c}
F'_x + F'_z p + F'_p \dfrac{dp}{dx} + F'_q \dfrac{dq}{dx} = 0, & \Phi'_p, \\
\Phi'_x + \Phi'_z p + \Phi'_p \dfrac{dp}{dx} + \Phi'_q \dfrac{dq}{dx} = 0, & -F'_p, \\
F'_y + F'_z q + F'_p \dfrac{dp}{dy} + F'_q \dfrac{dq}{dy} = 0, & \Phi'_q, \\
\Phi'_y + \Phi'_z q + \Phi'_p \dfrac{dp}{dy} + \Phi'_q \dfrac{dq}{dy} = 0, & -F'_q, \\
\dfrac{dp}{dy} - \dfrac{dq}{dx} = 0, & F'_q\Phi'_p - \Phi'_q F'_p.
\end{array}$$

Multiply by the quantities on the right and add. Then

$$(F'_x + pF'_z)\frac{\partial\Phi}{\partial p} + (F'_y + qF'_z)\frac{\partial\Phi}{\partial q} - F'_p\frac{\partial\Phi}{\partial x} - F'_q\frac{\partial\Phi}{\partial y} - (pF'_p + qF'_q)\frac{\partial\Phi}{\partial z} = 0. \tag{20}$$

Now this is a linear equation for Φ and is equivalent to

$$\frac{dp}{F'_x + pF'_z} = \frac{dq}{F'_y + qF'_z} = \frac{dx}{-F'_p} = \frac{dy}{-F'_q} = \frac{dz}{-(pF'_p + qF'_q)} = \frac{d\Phi}{0}. \tag{21}$$

Any integral of this system containing p or q and a will do for Φ, and the simplest integral will naturally be chosen.

As an example take $zp(x + y) + p(q - p) - z^2 = 0$. Then Charpit's equations are

$$\frac{dp}{-zp + p^2(x+y)} = \frac{dq}{zp - 2zq + pq(x+y)} = \frac{dx}{2p - q - z(x+y)}$$
$$= \frac{dy}{-p} = \frac{dz}{2p^2 - 2pq - pz(x+y)}.$$

How to combine these so as to get a solution is not very clear. Suppose the substitution $z = e^{z'}$, $p = e^{z'}p'$, $q = e^{z'}q'$ be made in the equation. Then

$$p'(x + y) + p'(q' - p') - 1 = 0$$

is the new equation. For this Charpit's simultaneous system is

$$\frac{dp'}{p'} = \frac{dq'}{p'} = \frac{dx}{2p' - q' - (x+y)} = \frac{dy}{-p'} = \frac{dz}{2p'^2 - 2pq - p'(x+y)}.$$

The first two equations give at once the solution $dp' = dq'$ or $q' = p' + a$. Solving

$$p'(x + y) + p'(q' - p') - 1 = 0 \quad \text{and} \quad q' = p' + a,$$

$$p' = \frac{1}{a + x + y}, \qquad q' = \frac{1}{a + x + y} + a, \qquad dz' = \frac{dx + dy}{a + x + y} + a\,dy.$$

Then $z = \log(a+x+y) + ay + b$ or $\log z = \log(a+x+y) + ay + b$

is a complete solution of the given equation. This will determine the general integral by eliminating a between the three equations

$$z = e^{ay+b}(a+x+y), \qquad b = f(a), \qquad 0 = (y + f'(a))(a+x+y) + 1,$$

where $f(a)$ denotes an arbitrary function. The rules for determining the singular solution give $z = 0$; but it is clear that the surfaces in the complete solution cannot be tangent to the plane $z = 0$ and hence the result $z = 0$ must be not a singular solution but an extraneous factor. There is no singular solution.

The method of solving a partial differential equation of higher order than the first is to reduce it first to an equation of the first order and then to complete the integration. Frequently the form of the equation will suggest some method easily applied. For instance, if the derivatives of lower order corresponding to one of the independent variables are absent, an integration may be performed as if the equation were an ordinary equation with that variable constant, and the constant of integration may be taken as a function of that variable. Sometimes a change of variable or an interchange of one of the independent variables with the dependent variable will simplify the equation. In general the solver is left mainly to his own devices. Two special methods will be mentioned below.

117. If the equation is *linear with constant coefficients* and all the derivatives are of the same order, the equation is

$$(a_0D_x^n + a_1D_x^{n-1}D_y + \cdots + a_{n-1}D_xD_y^{n-1} + a_nD_y^n)z = R(x, y). \qquad (22)$$

Methods like those of § 95 may be applied. Factor the equation.

$$a_0(D_x - \alpha_1D_y)(D_x - \alpha_2D_y)\cdots(D_x - \alpha_nD_y)z = R(x, y). \qquad (22')$$

Then the equation is reduced to a succession of equations

$$D_xz - \alpha D_yz = R(x, y),$$

each of which is linear of the first order (and with constant coefficients). Short cuts analogous to those previously given may be developed, but will not be given. If the derivatives are not all of the same order but the polynomial can be factored into linear factors, the same method will apply. For those interested, the several exercises given below will serve as a synopsis for dealing with these types of equation.

There is one equation of the second order,* namely

$$\frac{1}{V^2}\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2}, \qquad (23)$$

* This is one of the important differential equations of physics; other important equations and methods of treating them are discussed in Chap. XX.

which occurs constantly in the discussion of waves and which has therefore the name of the *wave equation.* The solution may be written down by inspection. For try the form

$$u(x, y, z, t) = F(ax + by + cz - Vt) + G(ax + by + cz + Vt). \quad (24)$$

Substitution in the equation shows that this is a solution if the relation $a^2 + b^2 + c^2 = 1$ holds, no matter what functions F and G may be. Note that the equation

$$ax + by + cz - Vt = 0, \qquad a^2 + b^2 + c^2 = 1,$$

is the equation of a plane at a perpendicular distance Vt from the origin along the direction whose cosines are a, b, c. If t denotes the time and if the plane moves away from the origin with a velocity V, the function $F(ax + by + cz - Vt) = F(0)$ remains constant; and if $G = 0$, the value of u will remain constant. Thus $u = F$ represents a phenomenon which is constant over a plane and retreats with a velocity V, that is, a plane wave. In a similar manner $u = G$ represents a plane wave approaching the origin. The general solution of (23) therefore represents the superposition of an advancing and a retreating plane wave.

To Monge is due a method sometimes useful in treating differential equations of the second order linear in the derivatives r, s, t; it is known as *Monge's method.* Let

$$Rr + Ss + Tt = V \quad (25)$$

be the equation, where R, S, T, V are functions of the variables and the derivatives p and q. From the given equation and

$$dp = rdx + sdy, \qquad dq = sdx + tdy,$$

the elimination of r and t gives the equation

$$s(Rdy^2 - Sdxdy + Tdx^2) - (Rdydp + Tdxdq - Vdxdy) = 0,$$

and this will surely be satisfied if the two equations

$$Rdy^2 - Sdxdy + Tdx^2 = 0, \qquad Rdydp + Tdxdq - Vdxdy = 0 \quad (25')$$

can be satisfied simultaneously. The first may be factored as

$$dy - f_1(x, y, z, p, q)dx = 0, \qquad dy - f_2(x, y, z, p, q)dx = 0. \quad (26)$$

The problem then is reduced to integrating the system consisting of one of these factors with (25′) and $dz = pdx + qdy$, that is, a system of three total differential equations.

If two independent solutions of this system can be found, as

$$u_1(x, y, z, p, q) = C_1, \qquad u_2(x, y, z, p, q) = C_2,$$

then $u_1 = \Phi(u_2)$ is a first or intermediary integral of the given equation, the general integral of which may be found by integrating this equation of the first order. If the two factors are distinct, it may happen that the two systems which arise may both be integrated. Then two first integrals $u_1 = \Phi(u_2)$ and $v_1 = \Psi(v_2)$ will be found, and instead of integrating one of these equations it may be better to solve both for p and q and to substitute in the expression $dz = pdx + qdy$ and integrate. When, however, it is not possible to find even one first integral, Monge's method fails.

As an example take $(x+y)(r-t)=-4p$. The equations are

$$(x+y)\,dy^2-(x+y)\,dx^2=0 \quad \text{or} \quad dy-dx=0, \qquad dy+dx=0$$

and
$$(x+y)\,dy\,dp-(x+y)\,dx\,dq+4p\,dx\,dy=0. \tag{A}$$

Now the equation $dy-dx=0$ may be integrated at once to give $y=x+C_1$. The second equation (A) then takes the form

$$2x\,dp+4p\,dx-2x\,dq+C_1(dp-dq)=0;$$

but as $dz=p\,dx+q\,dy=(p+q)\,dx$ in this case, we have by combination

$$2(x\,dp+p\,dx)-2(x\,dq+q\,dx)+C_1(dp-dq)+2\,dz=0$$

or
$$(2x+C_1)(p-q)+2z=C_2 \quad \text{or} \quad (x+y)(p-q)+2z=C_2.$$

Hence
$$(x+y)(p-q)+2z=\Phi(y-x) \tag{27}$$

is a first integral. This is linear and may be integrated by

$$\frac{dx}{x+y}=-\frac{dy}{x+y}=\frac{dz}{\Phi(y-x)-2z} \quad \text{or} \quad x+y=K_1, \qquad \frac{dx}{K_1}=\frac{dz}{\Phi(K_1-2x)-2z}.$$

This equation is an ordinary linear equation in z and x. The integration gives

$$K_1 z e^{\frac{2x}{K_1}}=\int e^{\frac{2x}{K_1}}\Phi(K_1-2x)\,dx+K_2.$$

Hence
$$(x+y)\,z e^{\frac{2x}{x+y}}-\int e^{\frac{2x}{K_1}}\Phi(K_1-2x)\,dx=K_2=\Psi(K_1)=\Psi(x+y)$$

is the general integral of the given equation when K_1 has been replaced by $x+y$ after integration, — an integration which cannot be performed until Φ is given.

The other method of solution would be to use also the second system containing $dy+dx=0$ instead of $dy-dx=0$. Thus in addition to the first integral (27) a second intermediary integral might be sought. The substitution of $dy+dx=0$, $y+x=C_1$ in (A) gives $C_1(dp+dq)+4p\,dx=0$. This equation is not integrable, because $dp+dq$ is a perfect differential and $p\,dx$ is not. The combination with $dz=p\,dx+q\,dy=(p-q)\,dx$ does not improve matters. Hence it is impossible to determine a second intermediary integral, and the method of completing the solution by integrating (27) is the only available method.

Take the equation $ps-qr=0$. Here $S=p$, $R=-q$, $T=V=0$. Then

$$-q\,dy^2-p\,dx\,dy=0 \quad \text{or} \quad dy=0, \quad p\,dx+q\,dy=0 \quad \text{and} \quad -q\,dy\,dp=0$$

are the equations to work with. The system $dy=0$, $q\,dy\,dp=0$, $dz=p\,dx+q\,dy$, and the system $p\,dx+q\,dy=0$, $q\,dy\,dp=0$, $dz=p\,dx+q\,dy$ are not very satisfactory for obtaining an intermediary integral $u_1=\Phi(u_2)$, although $p=\Phi(z)$ is an obvious solution of the first set. It is better to use a method adapted to this special equation. Note that

$$\frac{\partial}{\partial x}\left(\frac{q}{p}\right)=\frac{ps-qr}{p^2}, \quad \text{and} \quad \frac{\partial}{\partial x}\left(\frac{q}{p}\right)=0 \quad \text{gives} \quad \frac{q}{p}=f(y).$$

By (11), p. 124,
$$\frac{q}{p}=-\left(\frac{\partial x}{\partial y}\right)_z; \quad \text{then} \quad \frac{\partial x}{\partial y}=-f(y)$$

and
$$x=-\int f(y)\,dy+\Psi(z)=\Phi(y)+\Psi(z).$$

EXERCISES

1. Integrate these equations and discuss the singular solution:

(α) $p^{\frac{1}{2}} + q^{\frac{1}{2}} = 2x$, (β) $(p^2 + q^2)x = pz$, (γ) $(p + q)(px + qy) = 1$,
(δ) $pq = px + qy$, (ϵ) $p^2 + q^2 = x + y$, (ζ) $xp^2 - 2zp + xy = 0$,
(η) $q^2 = z^2(p - q)$, (θ) $q(p^2z + q^2) = 1$, (ι) $p(1 + q^2) = q(z - c)$,
(κ) $xp(1 + q) = qz$, (λ) $y^2(p^2 - 1) = x^2p^2$, (μ) $z^2(p^2 + q^2 + 1) = c^2$,
(ν) $p = (z + yq)^2$, (o) $pz = 1 + q^2$, (π) $z - pq = 0$, (ρ) $q = xp + p^2$.

2. Show that the rule for the type of Ex. 13, p. 273, can be deduced by Charpit's method. How about the generalized Clairaut form of Ex. 15?

3. (α) For the solution of the type $f_1(x, p) = f_2(y, q)$, the rule is: Set

$$f_1(x, p) = f_2(y, q) = a,$$

and solve for p and q as $p = g_1(x, a)$, $q = g_2(y, a)$; the complete solution is

$$z = \int g_1(x, a)\,dx + \int g_2(y, a)\,dy + b.$$

(β) For the type $F(z, p, q) = 0$ the rule is: Set $X = x + ay$, solve

$$F\left(z, \frac{dz}{dX}, a\frac{dz}{dX}\right) \text{ for } \frac{dz}{dX} = \phi(z, a), \text{ and let } \int\frac{dz}{\phi(z, a)} = f(z, a);$$

the complete solution is $x + ay + b = f(z, a)$. Discuss these rules in the light of Charpit's method. Establish a rule for the type $F(x + y, p, q) = 0$. Is there any advantage in using the rules over the use of the general method? Assort the examples of Ex. 1 according to these rules as far as possible.

4. What is obtainable for partial differential equations out of any characteristics of homogeneity that may be present?

5. By differentiating $p = f(x, y, z, q)$ successively with respect to x and y show that the expansion of the solution by Taylor's Formula about the point (x_0, y_0, z_0) may be found if the successive derivatives with respect to y alone,

$$\frac{\partial z}{\partial y}, \quad \frac{\partial^2 z}{\partial y^2}, \quad \frac{\partial^3 z}{\partial y^3}, \quad \ldots, \quad \frac{\partial^n z}{\partial y^n}, \quad \ldots,$$

are assigned arbitrary values at that point. Note that this arbitrariness allows the solution to be passed through any curve through (x_0, y_0, z_0) in the plane $x = x_0$.

6. Show that $F(x, y, z, p, q) = 0$ satisfies Charpit's equations

$$du = \frac{dx}{-F'_p} = \frac{dy}{-F'_q} = \frac{dz}{-(pF'_p + qF'_q)} = \frac{dp}{F'_x + pF'_z} = \frac{dq}{F'_y + qF'_z}, \tag{28}$$

where u is an auxiliary variable introduced for symmetry. Show that the first three equations are the differential equations of the lineal elements of the cones of Ex. 6, p. 272. The integrals of (28) therefore define a system of curves which have a planar element of the equation $F = 0$ passing through each of their lineal tangential elements. If the equations be integrated and the results be solved for the variables, and if the constants be so determined as to specify one particular curve with the initial conditions x_0, y_0, z_0, p_0, q_0, then

$$x = x(u, x_0, y_0, z_0, p_0, q_0), \quad y = y(\cdots), z = z(\cdots), \quad p = p(\cdots), \quad q = q(\cdots).$$

Note that, along the curve, $q = f(p)$ and that consequently the planar elements just mentioned must lie upon a developable surface containing the curve (§ 67). The curve and the planar elements along it are called a characteristic and a *characteristic strip* of the given differential equation. In the case of the linear equation the characteristic curves afforded the integration and any planar element through their lineal tangential elements satisfied the equation; but here it is only those planar elements which constitute the characteristic strip that satisfy the equation. What the complete integral does is to piece the characteristic strips into a family of surfaces dependent on two parameters.

7. By simple devices integrate the equations. Check the answers:

$(\alpha)\ \dfrac{\partial^2 z}{\partial x^2} = f(x)$, $\quad(\beta)\ \dfrac{\partial^n z}{\partial y^n} = 0$, $\quad(\gamma)\ \dfrac{\partial^2 z}{\partial x \partial y} = \dfrac{x}{y} + a$,

$(\delta)\ s + pf(x) = g(y)$, $\quad(\epsilon)\ ar = xy$, $\quad(\zeta)\ xr = (n-1)p$.

8. Integrate these equations by the method of factoring:

$(\alpha)\ (D_x^2 - a^2 D_y^2) z = 0$, $\quad(\beta)\ (D_x - D_y)^3 z = 0$, $\quad(\gamma)\ (D_x D_y^2 - D_y^3) z = 0$,

$(\delta)\ (D_x^2 + 3 D_x D_y + 2 D_y^2) z = x + y$, $\quad(\epsilon)\ (D_x^2 - D_x D_y - 6 D_y^2) z = xy$,

$(\zeta)\ (D_x^2 - D_y^2 - 3 D_x + 3 D_y) z = 0$, $\quad(\eta)\ (D_x^2 - D_y^2 + 2 D_x + 1) z = e^{-x}$.

9. Prove the operational equations:

$(\alpha)\ e^{axD_y} \phi(y) = (1 + axD_y + \frac{1}{2} a^2 x^2 D_y^2 + \cdots) \phi(y) = \phi(y + ax)$,

$(\beta)\ \dfrac{1}{D_x - aD_y} 0 = e^{axD_y} \dfrac{1}{D_x} 0 = e^{axD_y} \phi(y) = \phi(y + ax)$,

$(\gamma)\ \dfrac{1}{D_x - aD_y} R(x, y) = e^{axD_y} \displaystyle\int^x e^{-a\xi D_y} R(\xi, y)\, d\xi = \int^x R(\xi, y + ax - a\xi)\, d\xi$.

10. Prove that if $[(D_x - \alpha_1 D_y)^{m_1} \cdots (D_x - \alpha_k D_y)^{m_k}] z = 0$, then

$$z = \Phi_{11}(y + \alpha_1 x) + x\Phi_{12}(y + \alpha_1 x) + \cdots + x^{m_1 - 1}\Phi_{1m_1}(y + \alpha_1 x) + \cdots$$
$$+ \Phi_{k1}(y + \alpha_k x) + x\Phi_{k2}(y + \alpha_k x) + \cdots + x^{m_k - 1}\Phi_{km_k}(y + \alpha_k x),$$

where the Φ's are all arbitrary functions. This gives the solution of the reduced equation in the simplest case. What terms would correspond to $(D_x - \alpha D_y - \beta)^m z = 0$?

11. Write the solutions of the equations (or equations reduced) of Ex. 8.

12. State the rule of Ex. 9 (γ) as: Integrate $R(x, y - ax)$ with respect to x and in the result change y to $y + ax$. Apply this to obtaining particular solutions of Ex. 8 (δ), (ϵ), (η) with the aid of any short cuts that are analogous to those of Chap. VIII.

13. Integrate the following equations:

$(\alpha)\ (D_x^2 - D_{xy}^2 + D_y - 1) z = \cos(x + 2y) + e^y$, $\quad(\beta)\ x^2 r^2 + 2xys + y^2 t^2 = x^2 + y^2$,

$(\gamma)\ (D_x^2 + D_{xy} + D_y - 1) z = \sin(x + 2y)$, $\quad(\delta)\ r - t - 3p + 3q = e^{x+2y}$,

$(\epsilon)\ (D_x^3 - 2 D_x D_y^2 + D_y^3) z = x^{-2}$, $\quad(\zeta)\ r - t + p + 3q - 2z = e^{x-y} - x^2 y$,

$(\eta)\ (D_x^2 - D_x D_y - 2 D_y^2 + 2 D_x + 2 D_y) z = e^{2x+3y} + \sin(2x + y) + xy$.

14. Try Monge's method on these equations of the second order:

$(\alpha)\ q^2 r - 2pqs + p^2 t = 0$, $\quad(\beta)\ r - a^2 t = 0$, $\quad(\gamma)\ r + s = -p$,

$(\delta)\ q(1 + q) r - (p + q + 2pq) s + p(1 + p) t = 0$, $\quad(\epsilon)\ x^2 r + 2xys + y^2 t = 0$,

$(\zeta)\ (b + cq)^2 r - 2(b + cq)(a + cp) s + (a + cp)^2 t = 0$, $\quad(\eta)\ r + ka^2 t = 2as$.

If any simpler method is available, state what it is and apply it also.

15. Show that an equation of the form $Rr + Ss + Tt + U(rt - s^2) = V$ necessarily arises from the elimination of the arbitrary function from

$$u_1(x, y, z, p, q) = f[u_2(x, y, z, p, q)].$$

Note that only such an equation can have an intermediary integral.

16. Treat the more general equation of Ex. 15 by the methods of the text and thus show that an intermediary integral may be sought by solving one of the systems

$$\begin{array}{ll} Udy + \lambda_1 Tdx + \lambda_1 Udp = 0, & Udx + \lambda_1 Rdy + \lambda_1 Udq = 0, \\ Udx + \lambda_2 Rdy + \lambda_2 Udq = 0, & Udy + \lambda_2 Tdx + \lambda_2 Udp = 0, \\ dz = pdx + qdy, & dz = pdx + qdy, \end{array}$$

where λ_1 and λ_2 are roots of the equation $\lambda^2(RT + UV) + \lambda US + U^2 = 0$.

17. Solve the equations: (α) $s^2 - rt = 0$, (β) $s^2 - rt = a^2$,
(γ) $ar + bs + ct + e(rt - s^2) = h$, (δ) $xqr + ypt + xy(s^2 - rt) = pq$.