THE INTEGRATION OF FUNCTIONS OF A SINGLE VARIABLE

I. Introduction

The problem considered in the following pages is what is sometimes called the problem of 'indefinite integration' or of 'finding a function whose differential coefficient is a given function'. These descriptions are vague and in some ways misleading; and it is necessary to define our problem more precisely before we proceed further.

Let us suppose for the moment that $f(x)$ is a real continuous function of the real variable x. We wish to determine a function y whose differential coefficient is $f(x)$, or to solve the equation

$$\frac{dy}{dx}=f(x) \qquad \text{.............................(1).}$$

A little reflection shows that this problem may be analysed into a number of parts.

We wish, first, to know whether such a function as y necessarily exists, whether the equation (1) has always a solution; whether the solution, if it exists, is unique; and what relations hold between different solutions, if there are more than one. The answers to these questions are contained in that part of the theory of functions of a real variable which deals with 'definite integrals'. The definite integral

$$y=\int_a^x f(t)\,dt \qquad \text{.........................(2),}$$

which is defined as the limit of a certain sum, is a solution of the equation (1). Further

$$y+C \text{ ....(3),}$$

where C is an arbitrary constant, is also a solution, and all solutions of (1) are of the form (3).

These results we shall take for granted. The questions with which we shall be concerned are of a quite different character. They are questions as to the functional form of y when $f(x)$ is a function of some stated form. It is sometimes said that the problem of indefinite integration is that of 'finding an actual expression for y when $f(x)$ is given'. This statement is however still lacking in precision. The theory of definite integrals provides us not only with a proof of the existence of a solution, but also with an expression for it, an expression in the form of a limit. The problem of indefinite integration can be stated precisely only when we introduce sweeping restrictions as to the classes of functions and the modes of expression which we are considering.

Let us suppose that $f(x)$ belongs to some special class of functions $\mathfrak{F}$. Then we may ask whether y is itself a member of $\mathfrak{F}$, or can be expressed, according to some simple standard mode of expression, in terms of functions which are members of $\mathfrak{F}$. To take a trivial example, we might suppose that $\mathfrak{F}$ is the class of polynomials with rational coefficients: the answer would then be that y is in all cases itself a member of $\mathfrak{F}$.

The range and difficulty of our problem will depend upon our choice of (1) a class of functions and (2) a standard 'mode of expression'. We shall, for the purposes of this tract, take $\mathfrak{F}$ to be the class of *elementary functions*, a class which will be defined precisely in the next section, and our mode of expression to be that of *explicit expression in finite terms*, *i.e.* by formulae which do not involve passages to a limit.

One or two more preliminary remarks are needed. The subject-matter of the tract forms a chapter in the 'integral calculus'*, but does not depend in any way on any direct theory of integration. Such an equation as

$$y = \int f(x)\,dx \qquad \text{.............................(4)}$$

is to be regarded as merely another way of writing (1): the integral sign is used merely on grounds of technical convenience, and might be eliminated throughout without any substantial change in the argument.

* Euler, the first systematic writer on the 'integral calculus', defined it in a manner which identifies it with the theory of differential equations: 'calculus integralis est methodus, ex data differentialium relatione inveniendi relationem ipsarum quantitatum' (*Institutiones calculi integralis*, p. 1). We are concerned only with the special equation (1), but all the remarks we have made may be generalised so as to apply to the wider theory.

The variable x is in general supposed to be complex. But the tract should be intelligible to a reader who is not acquainted with the theory of analytic functions and who regards x as real and the functions of x which occur as real or complex functions of a real variable.

The functions with which we shall be dealing will always be such as are regular except for certain special values of x. These values of x we shall simply ignore. The meaning of such an equation as

$$\int \frac{dx}{x} = \log x$$

is in no way affected by the fact that $1/x$ and $\log x$ have infinities for $x = 0$.

II. Elementary functions and their classification

An *elementary function* is a member of the class of functions which comprises

(i) rational functions,

(ii) algebraical functions, explicit or implicit,

(iii) the exponential function e^x,

(iv) the logarithmic function $\log x$,

(v) all functions which can be defined by means of any finite combination of the symbols proper to the preceding four classes of functions.

A few remarks and examples may help to elucidate this definition.

1. A *rational function* is a function defined by means of any finite combination of the elementary operations of addition, multiplication, and division, operating on the variable x.

It is shown in elementary algebra that any rational function of x may be expressed in the form

$$f(x) = \frac{a_0 x^m + a_1 x^{m-1} + \dots + a_m}{b_0 x^n + b_1 x^{n-1} + \dots + b_n},$$

where m and n are positive integers, the a's and b's are constants, and the numerator and denominator have no common factor. We shall adopt this expression as the standard form of a rational function. It is hardly necessary to remark that it is in no way involved in the definition of a rational function that these constants should be rational or algebraical* or real *numbers*. Thus

$$\frac{x^2 + x + i\sqrt{2}}{x\sqrt{2} - e}$$

is a rational function.

2. An *explicit algebraical function* is a function defined by means of any finite combination of the four elementary operations and any finite number of operations of root extraction. Thus

$$\frac{\sqrt{(1+x)} - \sqrt[3]{(1-x)}}{\sqrt{(1+x)} + \sqrt[3]{(1-x)}},\quad \sqrt{\{x + \sqrt{(x + \sqrt{x})}\}},\quad \left(\frac{x^2 + x + i\sqrt{2}}{x\sqrt{2} - e}\right)^{\frac{2}{3}}$$

are explicit algebraical functions. And so is $x^{m/n}$ (*i.e.* $\sqrt[n]{x^m}$) for any integral values of m and n. On the other hand

$$x^{\sqrt{2}},\quad x^{1+i}$$

are not algebraical functions at all, but transcendental functions, as irrational or complex powers are defined by the aid of exponentials and logarithms.

Any explicit algebraical function of x satisfies an equation

$$P_0 y^n + P_1 y^{n-1} + \dots + P_n = 0$$

whose coefficients are polynomials in x. Thus, for example, the function

$$y = \sqrt{x} + \sqrt{(x + \sqrt{x})}$$

satisfies the equation

$$y^4 - (4y^2 + 4y + 1)x = 0.$$

The converse is not true, since it has been proved that in general equations of degree higher than the fourth have no roots which are explicit algebraical functions of their coefficients. A simple example is given by the equation

$$y^5 - y - x = 0.$$

We are thus led to consider a more general class of functions, *implicit* algebraical functions, which includes the class of explicit algebraical functions.

* An algebraical number is a number which is the root of an algebraical equation whose coefficients are integral. It is known that there are numbers (such as e and π) which are not roots of any such equation. See, for example, Hobson's *Squaring the circle* (Cambridge, 1913).

3. An *algebraical function* of x is a function which satisfies an equation

$$P_0 y^n + P_1 y^{n-1} + \dots + P_n = 0 \quad \text{..................(1)}$$

whose coefficients are polynomials in x.

Let us denote by $P(x, y)$ a polynomial such as occurs on the left-hand side of (1). Then there are two possibilities as regards any particular polynomial $P(x, y)$. Either it is possible to express $P(x, y)$ as the product of two polynomials of the same type, neither of which is a mere constant, or it is not. In the first case $P(x, y)$ is said to be *reducible*, in the second *irreducible*. Thus

$$y^4 - x^2 = (y^2 + x)(y^2 - x)$$

is reducible, while both $y^2 + x$ and $y^2 - x$ are irreducible.

The equation (1) is said to be reducible or irreducible according as its left-hand side is reducible or irreducible. A reducible equation can always be replaced by the logical alternative of a number of irreducible equations. Reducible equations are therefore of subsidiary importance only; and we shall always suppose that the equation (1) is irreducible.

An algebraical function of x is regular except at a finite number of points which are *poles* or *branch points* of the function. Let D be any closed simply connected domain in the plane of x which does not include any branch point. Then there are n and only n distinct functions which are one-valued in D and satisfy the equation (1). These n functions will be called the *roots* of (1) in D. Thus if we write

$$x = r(\cos\theta + i\sin\theta),$$

where $-\pi < \theta \leqslant \pi$, then the roots of

$$y^2 - x = 0,$$

in the domain

$$0 < r_1 \leqslant r \leqslant r_2, \quad -\pi < -\pi + \delta \leqslant \theta \leqslant \pi - \delta < \pi,$$

are $\sqrt{x}$ and $-\sqrt{x}$, where

$$\sqrt{x} = \sqrt{r}(\cos\tfrac{1}{2}\theta + i\sin\tfrac{1}{2}\theta).$$

The relations which hold between the different roots of (1) are of the greatest importance in the theory of functions*. For our present purposes we require only the two which follow.

(i) Any symmetric polynomial in the roots $y_1, y_2, \dots, y_n$ of (1) is a rational function of x.

* For fuller information the reader may be referred to Appell and Goursat's *Théorie des fonctions algébriques*.

(ii) Any symmetric polynomial in $y_2, y_3, \ldots, y_n$ is a polynomial in y_1 with coefficients which are rational functions of x.

The first proposition follows directly from the equations

$$\Sigma y_1 y_2 \ldots y_s = (-1)^s (P_{n-s}/P_0) \qquad (s = 1, 2, \ldots, n).$$

To prove the second we observe that

$$\sum_{2,3,\ldots} y_2 y_3 \ldots y_s = \sum_{1,2,\ldots} y_1 y_2 \ldots y_{s-1} - y_1 \sum_{2,3,\ldots} y_2 y_3 \ldots y_{s-1}.$$

so that the theorem is true for $\Sigma y_2 y_3 \ldots y_s$ if it is true for $\Sigma y_2 y_3 \ldots y_{s-1}$. It is certainly true for

$$y_2 + y_3 + \ldots + y_n = (y_1 + y_2 + \ldots + y_n) - y_1.$$

It is therefore true for $\Sigma y_2 y_3 \ldots y_s$, and so for any symmetric polynomial in $y_2, y_3, \ldots, y_n$.

4. Elementary functions which are not rational or algebraical are called *elementary transcendental functions* or elementary transcendents. They include all the remaining functions which are of ordinary occurrence in elementary analysis.

The trigonometrical (or circular) and hyperbolic functions, direct and inverse, may all be expressed in terms of exponential or logarithmic functions by means of the ordinary formulae of elementary trigonometry. Thus, for example,

$$\sin x = \frac{e^{ix} - e^{-ix}}{2i}, \qquad \sinh x = \frac{e^x - e^{-x}}{2},$$

$$\operatorname{arc} \tan x = \frac{1}{2i} \log \left(\frac{1+ix}{1-ix}\right), \quad \operatorname{arg} \tanh x = \frac{1}{2} \log \left(\frac{1+x}{1-x}\right).$$

There was therefore no need to specify them particularly in our definition.

The elementary transcendents have been further classified in a manner first indicated by Liouville*. According to him a function is a transcendent *of the first order* if the signs of exponentiation or of the taking of logarithms which occur in the formula which defines it apply only to rational or algebraical functions. For example

$$xe^{-x^2}, \quad e^{x^2} + e^x \sqrt{(\log x)}$$

are of the first order; and so is

$$\operatorname{arc} \tan \frac{y}{\sqrt{(1 + x^4)}},$$

* 'Mémoire sur la classification des transcendantes, et sur l'impossibilité d'exprimer les racines de certaines équations en fonction finie explicite des coefficients', *Journal de mathématiques*, ser. 1, vol. 2, 1837, pp. 56–104; 'Suite du mémoire...', *ibid.* vol. 3, 1838, pp. 523–546.

where y is defined by the equation

$$y^5 - y - x = 0;$$

and so is the function y defined by the equation

$$y^5 - y - e^x \log x = 0.$$

An elementary transcendent *of the second order* is one defined by a formula in which the exponentiations and takings of logarithms are applied to rational or algebraical functions or to transcendents of the first order. This class of functions includes many of great interest and importance, of which the simplest are

$$e^{e^x}, \ \log \log x.$$

It also includes irrational and complex powers of x, since, *e.g.*,

$$x^{\sqrt{2}} = e^{\sqrt{2} \log x}, \quad x^{1+i} = e^{(1+i) \log x};$$

the function

$$x^x = e^{x \log x};$$

and the logarithms of the circular functions.

It is of course presupposed in the definition of a transcendent of the second kind that the function in question is incapable of expression as one of the first kind or as a rational or algebraical function. The function

$$e^{\log R(x)},$$

where $R(x)$ is rational, is not a transcendent of the second kind, since it can be expressed in the simpler form $R(x)$.

It is obvious that we can in this way proceed to define transcendents of the nth order for all values of n. Thus

$$\log \log \log x, \ \log \log \log \log x, \ \ldots\ldots$$

are of the third, fourth, ... orders.

Of course a similar classification of algebraical functions can be and has been made. Thus we may say that

$$\sqrt{x}, \ \sqrt{(x + \sqrt{x})}, \ \sqrt{\{x + \sqrt{(x + \sqrt{x})}\}}, \ \ldots\ldots$$

are algebraical functions of the first, second, third, ... orders. But the fact that there is a general theory of algebraical equations and therefore of *implicit* algebraical functions has deprived this classification of most of its importance. There is no such general theory of elementary transcendental equations*, and therefore we shall not

* The natural generalisations of the theory of algebraical equations are to be found in parts of the theory of differential equations. See Königsberger, 'Bemerkungen zu Liouville's Classificirung der Transcendenten', *Math. Annalen*, vol. 28, 1886, pp. 483–492.

rank as 'elementary' functions defined by transcendental equations such as

$$y = x \log y,$$

but incapable (as Liouville has shown that in this case y is incapable) of explicit expression in finite terms.

5. The preceding analysis of elementary transcendental functions rests on the following theorems:

(*a*) e^x is not an algebraical function of x;

(*b*) $\log x$ is not an algebraical function of x;

(*c*) $\log x$ is not expressible in finite terms by means of signs of exponentiation and of algebraical operations, explicit or implicit*;

(*d*) transcendental functions of the first, second, third, ... orders actually exist.

A proof of the first two theorems will be given later, but limitations of space will prevent us from giving detailed proofs of the third and fourth. Liouville has given interesting extensions of some of these theorems: he has proved, for example, that no equation of the form

$$Ae^{\alpha p} + Be^{\beta p} + \dots + Re^{\rho p} = S,$$

where $p, A, B, \dots, R, S$ are algebraical functions of x, and $\alpha, \beta, \dots, \rho$ different constants, can hold for all values of x.

III. The integration of elementary functions. Summary of results

In the following pages we shall be concerned exclusively with the problem of the integration of elementary functions. We shall endeavour to give as complete an account as the space at our disposal permits of the progress which has been made by mathematicians towards the solution of the two following problems:

(i) *if $f(x)$ is an elementary function, how can we determine whether its integral is also an elementary function?*

(ii) *if the integral is an elementary function, how can we find it?*

It would be unreasonable to expect complete answers to these questions. But sufficient has been done to give us a tolerably complete insight into the nature of the answers, and to ensure that it

* For example, $\log x$ cannot be equal to e^y, where y is an algebraical function of x.

shall not be difficult to find the complete answers in any particular case which is at all likely to occur in elementary analysis or in its applications.

It will probably be well for us at this point to summarise the principal results which have been obtained.

1. The integral of a rational function (IV.) is *always* an elementary function. It is either rational or the sum of a rational function and of a finite number of constant multiples of logarithms of rational functions (IV., 1).

If certain constants which are the roots of an algebraical equation are treated as known then the form of the integral can always be determined completely. But as the roots of such equations are not in general capable of explicit expression in finite terms, it is not in general possible to express the integral in an absolutely explicit form (IV.; 2, 3).

We can always determine, by means of a finite number of the elementary operations of addition, multiplication, and division, whether the integral is rational or not. If it is rational, we can determine it completely by means of such operations; if not, we can determine its rational part (IV.; 4, 5).

The solution of the problem in the case of rational functions may therefore be said to be complete; for the difficulty with regard to the explicit solution of algebraical equations is one not of inadequate knowledge but of proved impossibility (IV., 6).

2. The integral of an algebraical function (V.), explicit or implicit, may or may not be elementary.

If y is an algebraical function of x then the integral $\int y\,dx$, or, more generally, the integral

$$\int R(x, y)\,dx,$$

where R denotes a rational function, is, if an elementary function, either algebraical or the sum of an algebraical function and of a finite number of constant multiples of logarithms of algebraical functions. All algebraical functions which occur in the integral are *rational functions of x and y* (V.; 11–14, 18).

These theorems give a precise statement of a general principle enunciated by Laplace*: '*l'intégrale d'une fonction différentielle*

* *Théorie analytique des probabilités*, p. 7.

(algébrique) ne peut contenir d'autres quantités radicaux que celles qui entrent dans cette fonction'; and, we may add, cannot contain *exponentials* at all. Thus it is impossible that

$$\int \frac{dx}{\sqrt{(1+x^2)}}$$

should contain e^x or $\sqrt{(1-x)}$: the appearance of these functions in the integral could only be apparent, and they could be eliminated before differentiation. Laplace's principle really rests on the fact, of which it is easy enough to convince oneself by a little reflection and the consideration of a few particular cases (though to give a rigorous proof is of course quite another matter), that *differentiation will not eliminate exponentials or algebraical irrationalities.* Nor, we may add, will it eliminate logarithms except when they occur in the simple form

$$A \log \phi(x),$$

where A is a constant, and this is why logarithms can only occur in this form in the integrals of rational or algebraical functions.

We have thus a general knowledge of the form of the integral of an algebraical function y, when it is itself an elementary function. Whether this is so or not of course depends on the nature of the equation $f(x, y) = 0$ which defines y. If this equation, when interpreted as that of a curve in the plane (x, y), represents a *unicursal* curve, *i.e.* a curve which has the maximum number of double points possible for a curve of its degree, or whose *deficiency* is zero, then x and y can be expressed simultaneously as rational functions of a third variable t, and the integral can be reduced by a substitution to that of a rational function (v.; 2, 7–9). In this case, therefore, the integral is always an elementary function. But this condition, though sufficient, is not necessary. It is in general true that, when $f(x, y) = 0$ is not unicursal, the integral is not an elementary function but a new transcendent; and we are able to classify these transcendents according to the deficiency of the curve. If, for example, the deficiency is unity, then the integral is in general a transcendent of the kind known as *elliptic integrals*, whose characteristic is that they can be transformed into integrals containing no other irrationality than the square root of a polynomial of the third or fourth degree (v., 20). But there are infinitely many cases in which the integral can be expressed by algebraical functions and logarithms. Similarly there are infinitely many cases in which integrals associated with curves whose deficiency is greater