

I trust that I have now made clear my own feelings regarding the three questions raised at the outset; first, as to why divergent series have come into such prominence since the appearance of the early volumes of the encyclopedia, second, what has been done that really constitutes a vital advance and third, as to whether such series are at last upon a truly scientific basis. My only fear is that in attempting to couch the whole in very simple form I may have gone too far in this direction and thus violated a principle which, I believe it is said, the poet Browning always carefully observed; namely, of never using so simple a style that the intelligence of one's readers or hearers may be offended. But this is a rather treacherous principle, as most people discover in attempting to read Browning, so I may perhaps be pardoned if I have seemed to depart too far from it.

SOLUTIONS OF DIFFERENTIAL EQUATIONS AS FUNCTIONS OF THE CONSTANTS OF INTEGRATION.

BY PROFESSOR GILBERT AMES BLISS.

(Read before the American Mathematical Society December 29, 1917.)

THE purpose of this note is to prove the differentiability of the solutions of a system of differential equations with respect to the constants of integration by a method which seems more natural and simpler than those which have hitherto been published. Incidentally a restatement of the so-called "imbedding theorem" for differential equations is given, a theorem which is frequently applied in the calculus of variations, and which has been useful, and could be made still more so, in many other connections. It is analogous to the fundamental theorem for implicit functions in its statement that a solution of a system of differential equations given in advance is always a member of a continuous family of such solutions.

Let C be an arc

$$(C) \quad x = u(\tau), \quad \tau_1 \leq \tau \leq \tau_2,$$

for which the function u is single-valued and has a continuous first derivative. The neighborhood C_ϵ of this arc is the totality of points (τ, x) which satisfy the inequalities

$$(C_\epsilon) \quad \tau_1 \leq \tau \leq \tau_2, \quad |x - u(\tau)| \leq \epsilon.$$

In the differential equation

$$(1) \quad \frac{dx}{d\tau} = f(\tau, x)$$

the function f is supposed

(a) to be single-valued and continuous in C_ϵ ;^{*}

(b) to satisfy the Lipschitz condition

$$(2) \quad |f(\tau, x) - f(\tau, x')| \leq \kappa |x - x'|$$

whenever (τ, x) and (τ, x') are both in C_ϵ ;

(c) to be such that the equation (1) has the arc C as a solution.

By a solution of equation (1) is always meant an arc of the type of C having a continuous derivative and satisfying the equation. The theorem to be proved is then the following:

For every neighborhood C_ϵ of the arc C with the properties just described there exists a second neighborhood C_δ through every point (τ_0, x_0) of which passes one and but one solution of equation (1), defined and in C_ϵ on the whole interval $\tau_1 \tau_2$. The function

$$x = v(\tau, \tau_0, x_0)$$

representing these solutions is continuous and has a continuous derivative $\partial v / \partial \tau$ in the region R of points (τ, τ_0, x_0) satisfying the conditions

$$(R) \quad \tau_1 \leq \tau \leq \tau_2, \quad (\tau_0, x_0) \text{ interior to } C_\delta.$$

If $f(\tau, x)$ has continuous partial derivatives of the n -th order in C_ϵ , then v and $\partial v / \partial \tau$ also have continuous partial derivatives up to and including those of order n when (τ, τ_0, x_0) lies in the region R .

The existence and continuity of the function $v(\tau, \tau_0, x_0)$ have been established in various ways. For the sake of completeness a proof will be given here which is based upon the method of approximation of Picard, and which makes use of the sequence of functions $\{v_m\}$ defined by the following equations:

^{*} If only the continuity of f with respect to τ is presupposed, then with the help of (b) it is provable that f is continuous in τ and x together. The proof for boundary points of C_ϵ is less direct than for interior points.

$$\begin{aligned} |v_2 - v_1| &= \left| \int_{\tau_0}^{\tau} [f(\tau, v_1) - f(\tau, v_0)] d\tau \right|, \\ &\leq 2\epsilon \frac{\kappa |\tau - \tau_0|}{1!} \end{aligned}$$

since v_0 and v_1 both define curves in C_ϵ and hence differ by at most 2ϵ . An induction as before now gives

$$|v_{m+1} - v_m| \leq 2\epsilon \frac{\kappa^m |\tau - \tau_0|^m}{m!} \leq 2\epsilon \frac{\kappa^m \lambda^m}{m!},$$

showing that the series

$$(5) \quad v_1 + (v_2 - v_1) + \cdots + (v_{m+1} - v_m) + \cdots$$

converges uniformly.

From the last of the equations (3) it follows as usual that the limit function $v(\tau, \tau_0, x_0)$ satisfies the equation

$$v(\tau, \tau_0, x_0) = x_0 + \int_{\tau_0}^{\tau} f[\tau, v(\tau, \tau_0, x_0)] d\tau$$

and hence also the differential equation (1). If there were a second solution $w(t)$ of equation (1) through the point (τ_0, x_0) it would satisfy an equation similar to the last one, and consequently also the inequality

$$|w - v| \leq \left| \int_{\tau_0}^{\tau} \kappa |w - v| d\tau \right| \leq 2\epsilon \frac{\kappa |\tau - \tau_0|}{1!}$$

since 2ϵ is greater than $|w - v|$ on the interval $\tau_1\tau_2$. But successive application of this relation gives

$$|w - v| \leq 2\epsilon \frac{\kappa^m |\tau - \tau_0|^m}{m!},$$

which can be true only if v and w are identical, since the second member has the limit zero as m increases.

The proofs of the preceding paragraphs establish the existence, continuity, and uniqueness of the function $v(\tau, \tau_0, x_0)$, and also of its derivative $\partial v/\partial \tau$ since v satisfies (1). To prove that it has the further derivatives described in the theorem, suppose first that $f(\tau, x)$ has continuous first derivatives in C_ϵ . Then each of the functions v_m has a derivative

$$(6) \quad \frac{\partial v_m}{\partial x_0} = 1 + \int_{\tau_0}^{\tau} f_x(\tau, v_{m-1}) \frac{\partial v_{m-1}}{\partial x_0} d\tau$$

and it is proposed to show that the sequence $\{\partial v_m/\partial x_0\}$ converges uniformly for values (τ, τ_0, x_0) in the region R . A well-known theorem* concerning the differentiation of a series term by term, applied to the series (5), then establishes the fact that the limit $v(\tau, \tau_0, x_0)$ of the sequence $\{v_m\}$ has as its derivative the limit of the sequence $\{\partial v_m/\partial x_0\}$.

Consider the two sums

$$(7) \quad B = 1 + \int_{\tau_0}^{\tau} f_x d\tau + \cdots + \int_{\tau_0}^{\tau} f_x \cdots \int_{\tau_0}^{\tau} f_x d\tau + \cdots,$$

$$(8) \quad \frac{\partial v_m}{\partial x_0} = 1 + \int_{\tau_0}^{\tau} A_{m-1} d\tau + \cdots + \int_{\tau_0}^{\tau} A_{m-1} \cdots \int_{\tau_0}^{\tau} A_0 d\tau^m,$$

where $f_x = f_x(\tau, v)$ represents the first partial derivative of f with respect to x , and $A_n = f_x(\tau, v_n)$. The first sum is an infinite series, while the second has a finite number of terms found by successive application of formula (6). A simple inductive proof shows that each term of (8) approaches uniformly the corresponding term of (7) as a limit when m increases indefinitely, since the elements of the sequence $\{v_m\}$ approach uniformly the limit v . Further each term of both sums has absolute value less than the corresponding term of the series

$$1 + \frac{\kappa |\tau - \tau_0|}{1!} + \cdots + \frac{\kappa^n |\tau - \tau_0|^n}{n!} + \cdots$$

or of the series

$$(9) \quad 1 + \frac{\kappa \lambda}{1!} + \cdots + \frac{\kappa^n \lambda^n}{n!} + \cdots,$$

where κ is now the maximum of the absolute value of f_x in C_ϵ . It follows that an integer n can be taken so large that all the terms of the series (7) after the n th have a sum with absolute value less than $\epsilon/3$, and also so large that the same is true for the expression (8). Since the individual terms of (8) approach uniformly those of (7) it is also true that for sufficiently large values of m the sums of the first n terms of (8) and (7) differ by less than $\epsilon/3$. Hence for such values of m the difference between $\partial v_m/\partial x_0$ and the sum of the series (7) is less than ϵ , and the derivatives $\partial v_m/\partial x_0$ approach the sum of the series (7) uniformly for all values (τ, τ_0, x_0) in the region R . The function v has therefore a derivative $\partial v/\partial x_0$ given by the series

* Goursat-Hedrick, A Course in Mathematical Analysis, vol. 1, p. 365.

$$(10) \quad \frac{\partial v}{\partial x_0} = 1 + \int_{\tau_0}^{\tau} f_x(\tau, v) d\tau + \dots \\ + \int_{\tau_0}^{\tau} f_x(\tau, v) \dots \int_{\tau_0}^{\tau} f_x(\tau, v) d\tau^m + \dots,$$

which converges uniformly, since its terms are respectively less in absolute value than those of (9), and $\partial v/\partial x_0$ is therefore continuous in the region R .

Similar reasoning shows the existence and continuity of the derivative $\partial v/\partial \tau_0$. For from equations (3)

$$\frac{\partial v_m}{\partial \tau_0} = -f(\tau_0, x_0) + \int_{\tau_0}^{\tau} A_{m-1} \frac{\partial v_{m-1}}{\partial \tau_0} d\tau \\ = -f_0 - \int_{\tau_0}^{\tau} A_{m-1} f_0 d\tau - \dots \\ - \int_{\tau_0}^{\tau} A_{m-1} \dots \int_{\tau_0}^{\tau} A_1 f_0 d\tau^{m-1} - \int_{\tau_0}^{\tau} A_{m-1} \dots \int_{\tau_0}^{\tau} A_0 u'(\tau_0) d\tau^m,$$

where $f_0 = f(\tau_0, x_0)$. The comparison series analogous to (9) is ν times that series, where ν is the larger of the maxima of $|f(\tau_0, x_0)|$ and $|u'(\tau_0)|$ in R , and the value of the derivative sought is the continuous sum of the uniformly convergent series

$$(11) \quad \frac{\partial v}{\partial \tau_0} = -f(\tau_0, x_0) - \int_{\tau_0}^{\tau} f_x(\tau, v) f(\tau_0, x_0) d\tau - \dots \\ - \int_{\tau_0}^{\tau} f_x(\tau, v) \dots \int_{\tau_0}^{\tau} f_x(\tau, v) f_0(\tau_0, x_0) d\tau^m - \dots.$$

Finally the equation

$$(12) \quad \frac{\partial v}{\partial \tau} = f(\tau, v)$$

shows that $\partial v/\partial \tau$ is continuous and has continuous first partial derivatives in the region R , since f and v have this property in C_ϵ and R , respectively. This completes the proof of the theorem for the case $n = 1$. The corresponding results for an arbitrary value of n are readily established by induction when the theorem for $n = 1$ has been proved for a system of equations instead of a single one, as will be shown below. In concluding the simpler case it is important to note the relation

$$(13) \quad \frac{\partial v}{\partial \tau_0} = - \frac{\partial v}{\partial x_0} f(\tau_0, x_0),$$

which follows readily from (11) and (10).

The proofs of the preceding theorems for a system of differential equations are of the same character as those above, but the equations used in the proofs will be much more complicated unless notations are used for matrices and multipartite numbers. Peano seems to have been the first to define and apply the moduli of multipartite numbers and matrices in proving existence theorems for differential equations, the case which he considered being that of a linear system.* If x denotes the set of numbers $(x', x'', \dots, x^{(p)})$, and f the set $(f', f'', \dots, f^{(p)})$, the single equation (1) will represent the system

$$(14) \quad \frac{dx^{(i)}}{dt} = f^{(i)}(t, x', \dots, x^{(p)}) \quad (i = 1, \dots, p).$$

By replacing absolute values everywhere by moduli, the modulus of x for example being

$$(15) \quad \text{mod } x = \sqrt{x'^2 + \dots + x^{(p)2}},$$

and using simple properties of multipartite numbers and matrices,† the proofs for the system (14) can be carried through quite simply and the notations will be the same, equation for equation, as in the paragraphs above. To facilitate the re-interpretation of the equations the following table of the symbols used is given:

Positive integers: i, j, m, n, p ;

Scalars: $\delta, \epsilon, \kappa, \lambda, \mu, \nu, \rho, \sigma, \tau, e$;

p -partite numbers: $a, f, x, y, z, u, v, w, \partial v/\partial \tau_0, \partial v_m/\partial \tau_0$;

p -square matrices: $A, B, f_x = \partial f/\partial x, \partial v/\partial x_0, \partial v_m/\partial x_0$.

In order to carry through the proofs of the existence and continuity of the derivatives of the p functions v , some additional properties of matrices are needed besides those given in the papers cited above. The definition of the modulus of a matrix given below is equivalent to that of Peano but different in form. It is closely associated with the notion of a limited

* *Mathematische Annalen*, vol. 32 (1888), p. 450.

† See also my paper, "The solutions of differential equations of the first order as functions of their initial values," *Annals of Mathematics*, 2d ser., vol. 6 (1905), p. 58.

matrix used by Hilbert in his theory of quadratic forms in a denumerable infinity of variables, and for the special case here considered is identical with the modulus as defined by E. H. Moore for the purposes of his General Analysis. It is applied here only to square matrices, but would be equally effective for those which are of unequal dimensions, and in particular for a matrix of one row and p columns it gives precisely the value (15). The table of properties in the next paragraph but one contains besides those of Peano some further properties which are useful in the present paper.

Let A be a p -square matrix, and let $\text{mod } A$ be the maximum of the bilinear form $Ay \cdot z^*$ on the set of values of y and z which satisfy the equations $\text{mod } y = \text{mod } z = 1$. This maximum is attained since the set over which y and z range is closed, and from the elementary theory of maxima and minima the values of y and z which determine the maximum satisfy the linear equations

$$(16) \quad Ay = \rho z, \quad \bar{A}z = \sigma y,$$

from which also

$$(17) \quad \bar{A}Ay = \rho\sigma y, \quad \bar{A}Az = \rho\sigma z,$$

where ρ, σ are scalars and \bar{A} is the matrix formed from A by interchanging rows and columns. From the equations (16) one deduces readily

$$(18) \quad Ay \cdot z = \rho, \quad \bar{A}z \cdot y = Ay \cdot z = \sigma,$$

when $\text{mod } y = \text{mod } z = 1$, so that ρ and σ are both equal to the value of the modulus of A . Furthermore from the equations (16) and (18)

$$Ay \cdot Ay = \bar{A}Ay \cdot y = \rho^2,$$

which suggests that the modulus defined above may be identical with that of Peano, the value which he used being the square root of the maximum of $Ay \cdot Ay$ on the set of values y satisfying $\text{mod } y = 1$. By reasoning quite similar to that just used, and with the help of equations (17), this proves to be the case.

The modulus of a matrix has the following properties, which are readily provable:

* The expression $y \cdot z$ is the sum of the products of the elements of y and z ; Ay is the m -partite number whose elements are the dot-products of the rows of A by y . For the use of the dot in this connection see Gibbs-Wilson, Vector Analysis, p. 55.

- (1) $\text{mod } A = \text{maximum of } Ay \cdot z \text{ on the set of values } y, z \text{ satisfying } \text{mod } y = \text{mod } z = 1,$
 $= \text{maximum of } \sqrt{Ay \cdot Ay} \text{ on the set of values } y \text{ satisfying } \text{mod } y = 1;$
- (2) every element of A has absolute value $\leq \text{mod } A$;
- (3) $\text{mod } A < p^2\delta$ if each element of A has absolute value less than δ ;
- (4) $\text{mod } \kappa A = |\kappa| \text{mod } A$;
 $\text{mod } Ay \leq \text{mod } A \text{ mod } y$;
 $\text{mod } AB \leq \text{mod } A \text{ mod } B$;
- (5) $\text{mod } (A + B) \leq \text{mod } A + \text{mod } B$,
 $\text{mod } (A - B) \leq |\text{mod } A - \text{mod } B|$;
- (6) $\text{mod } A$ is a continuous function of the elements of A ;
- (7) $\text{mod } \int_{\tau_0}^{\tau_1} A d\tau \leq \left| \int_{\tau_0}^{\tau_1} \text{mod } A d\tau \right|$;
- (8) $|\lambda\lambda| < \text{mod } (dA/d\tau)$ for every derivative number λ of $\text{mod } A$.

In the statement (7) the elements of A are supposed to be integrable functions of τ on the interval $\tau_0 \leq \tau \leq \tau_1$, and in (8) they are supposed to have unique derivatives at the value of τ considered. The property (8) is not necessary for the applications of the matrix theory in the present paper.

The property (2) holds because, by choosing all elements of y and z to be zero except one element equal to $+1$ or -1 in each, the value $Ay \cdot z$ may be made equal to the absolute value of a selected one of the elements of A . In the expression $Ay \cdot z$ with $\text{mod } y = \text{mod } z = 1$ there are p^2 terms each consisting of an element of A multiplied by two numbers with absolute values ≤ 1 . Hence when each term of A has absolute value less than δ the maximum of $Ay \cdot z$ surely does not exceed $p^2\delta$. The proof of the first formula (4) is immediate. The second part follows since for every y

$$\sqrt{Ay \cdot Ay} = \sqrt{\frac{Ay \cdot Ay}{(\text{mod } y)^2}} \text{mod } y \leq \text{mod } A \text{ mod } y.$$

To prove the third part of (4), suppose that y, z are sets of values having $\text{mod } y = \text{mod } z = 1$ and giving $ABy \cdot z$ its maximum. Then

$$\text{mod } AB = ABy \cdot z = \frac{ABy \cdot z}{\sqrt{By \cdot By}} \sqrt{By \cdot By} \leq \text{mod } A \text{ mod } B.$$

If y and z are similarly the values giving $(A + B)y \cdot z$ its maximum, then

$$\begin{aligned} \text{mod } (A + B) &= (A + B)y \cdot z \\ &= Ay \cdot z + By \cdot z \leq \text{mod } A + \text{mod } B. \end{aligned}$$

The second part of (5) follows from the first since

$$\begin{aligned} \text{mod } A &\leq \text{mod } (A - B) + \text{mod } B, \\ \text{mod } B &\leq \text{mod } (B - A) + \text{mod } A. \end{aligned}$$

If $A + \Delta A$ is a matrix formed from A by giving its elements increments with absolute values less than δ , then from the second property (5), and (3),

$$|\text{mod } (A + \Delta A) - \text{mod } A| \leq \text{mod } \Delta A < p^2 \delta,$$

which proves the statement (6). The property (7) is readily proved by applying the first formula (5) to the sum whose limit is the definite integral on the left, and then taking the limit. Finally the inequality

$$\left| \frac{\text{mod } (A + \Delta A) - \text{mod } A}{\Delta \tau} \right| \leq \text{mod } \frac{\Delta A}{\Delta \tau}$$

is deducible with the help of the second property (5), and it justifies (8) since every derivative number of $\text{mod } A$ is the limit of the fraction on the left when $\Delta \tau$ approaches zero over a suitably selected sequence of values.

In the equation (6) for the system (14) the symbol $\partial v_m / \partial x_0$ represents the matrix of derivatives $\|\partial v_m^{(i)} / \partial x_0^{(j)}\|$ ($i, j = 1, \dots, p$) and the first term on the right is to be interpreted as the identity matrix I . A similar agreement holds for the first terms of (7), (8), and (10). The symbol f_x now represents the matrix of derivatives $\|\partial f^{(i)} / \partial x^{(j)}\|$. Hence in (6) and the three last mentioned equations every term is a p -square matrix. With the help of the matrix properties (7) and (4₂) above

$$\begin{aligned} \text{mod } \int_{\tau_0}^{\tau} A_{m-1} \cdots \int_{\tau_0}^{\tau} A_{m-n} d\tau^n \\ \leq \left| \int_{\tau_0}^{\tau} \text{mod } A_{m-1} \cdots \int_{\tau_0}^{\tau} \text{mod } A_{m-n} d\tau^n \right| \leq \frac{\kappa^n |\tau - \tau_0|^n}{n!}, \end{aligned}$$

where κ is the maximum of $\text{mod } f_x$ in the region C_ϵ . Hence the terms of the series (9) exceed the moduli of the respective

terms in (7) and (8). An integer n can now be taken so large that the modulus of the sum of the terms of (8) after the n th is less than $\epsilon/3$, and so that the same is true for the sum of the terms after the n th in the series (7). Since the elements of the sequence $\{v_m\}$ converge uniformly to the elements of v , the elements of the matrix represented by the first n terms of the sum (8) converge to the corresponding elements of the similar matrix from the series (7), and by an application of the matrix property (3) it is clear that the modulus of the difference of these matrices is less than $\epsilon/3$ when $m > n$ is sufficiently large. For such values m therefore

$$\text{mod} \left(\frac{\partial v_m}{\partial x_0} - B \right) < \epsilon,$$

and because of the matrix property (2) each element of $\partial v_m / \partial x_0$ converges uniformly, as m increases, to the respective element of B . Hence every one of the elements of v has a derivative with respect to every element of x_0 and the matrix $\partial v / \partial x_0$ of these derivatives is B . Finally the series of moduli of the terms of (10) converges uniformly, implying also, on account of property (2), the uniform convergence of every series of corresponding elements in the matrices which the terms of (10) represent, and showing therefore that every element of the matrix $\partial v / \partial x_0$ is a continuous function.

The proof of the existence and continuity of the p derivatives $\partial v / \partial \tau_0$ is so similar that it is unnecessary to exhibit it in detail.

The proofs which have been given above establish the properties of the solution $v(\tau, \tau_0, x_0)$ of a system of equations for the case $n = 1$ described in the theorem, and the proof can now be given for the general case by an induction. Suppose that the existence and continuity of the n th derivatives of v and $\partial v / \partial \tau$ have been proved to be a consequence of the assumption that f has continuous n th derivatives, and suppose further that the functions f have continuous derivatives in C_ϵ of order $n + 1$. The system of $2p + 1$ equations

$$(19) \quad \frac{d\tau_0}{d\tau} = 0, \quad \frac{dx_0}{d\tau} = 0, \quad \frac{dw}{d\tau} = f_x(\tau, v)w$$

in the variables τ, τ_0, x_0, w has the solution

$$(17) \quad \tau_0 = \text{const.}, \quad x_0 = \text{const.}, \quad w = \frac{\partial v}{\partial x_0} w_0$$

with the initial values τ_0, x_0, w_0 for $\tau = \tau_0$, as one readily sees by differentiating equation (12) for the elements of x_0 and from the fact that $\partial v/\partial x_0$ is the identity matrix when $\tau = \tau_0$. There is a neighborhood Γ_ϵ in which the second members of equations (19) have continuous n th derivatives, provided that (τ, τ_0, x_0) is in R , since v has this property and f has by hypothesis continuous derivatives of order $n + 1$. Hence there is also a region P for the solution Γ , analogous to R for C , in which the solution Γ has continuous n th derivatives. This shows that all of the elements of the matrix $\partial v/\partial x_0$ have continuous n th derivatives since w_0 is an arbitrary constant multipartite number. The same is true of the elements of the multipartite function $\partial v/\partial \tau_0$ from equations (13), and for the elements of $\partial v/\partial \tau$ from equations (12). Hence all of the derivatives of v of order $n + 1$ exist and are continuous as described in the theorem. That $\partial v/\partial \tau$ has the same property follows at once from the relations (12).

In conclusion it may be remarked that the theorems concerning the dependence of the solutions v upon parameters involved in the equations (14) require no proofs essentially different from those given above. A system of the form

$$\frac{dx}{d\tau} = f(\tau, x, a),$$

where a is a multipartite set of parameters, may in fact be replaced by the system

$$\frac{dx}{d\tau} = f(\tau, x, a), \quad \frac{da}{d\tau} = 0$$

in the variables τ, x, a . If the functions f have continuous derivatives of order n in τ, x, a , then the theorem proved above justifies the statement that the derivatives of order n for the solutions $x = v(\tau, \tau_0, x, a)$ and their derivatives v_τ will also exist and be continuous in the region analogous to R for these equations.