4-1-1940

The existence theorem of ordinary differential equations

Harris J. Dark

THE UNIVERSITY OF RICHMOND

THE EXISTENCE THEOREM

OF

ORDINARY DIFFERENTIAL EQUATIONS

A Thesis

Submitted to the Faculty of

The Graduate School of the University of Richmond

as partial fulfillment of the

requirements for the degree of

Master of Arts

DEPARTMENT OF MATHEMATICS

By

Harris J. Dark

1940

CONTENTS

ACKNOWLEDGEMENT

The author wishes to acknowledge his debt of appreciation to Dr. Robert E. Gaines for a new conception and understanding of original work in mathematics and also to Dr. Charles H. Wheeler, who suggested the subject for this thesis, and who made many kind and helpful suggestions in the direction of this investigation.

CHAPTER I

INTRODUCTORY

There are a great many devices for solving differential equations of certain special forms. But there is a large number of classes of differential equations that are not included in these special forms and cannot be integrated by quadratures or other purely elementary methods. When mathematicians were forced to abandon their cherished hope of finding a method for expressing the solution of every differential equation in terms of a finite number of known functions or their integrals, they turned their attention to the question of whether a differential equation in general had a solution at all, and, if so, of what nature.

This study resulted in the development of what is known as the Existence Theorem of Ordinary Differential Equations. This theorem states that for every ordinary differential equation of a fairly general type there exists a solution. The type of equations included in the theorem includes those that are usually encountered and used both in applied and pure mathematics. The theorem is no less important in the field of calculus than is the cardinal proposition in the theory of algebraic equations, that every such equation has a root.

The theorem may be stated as follows:

Given a system of ordinary differential equations:

$$dy_1/dx = f_1(x, y_1, y_2, \ldots, y_m),$$

$$dy_2/dx = f_2(x, y_1, y_2, \ldots, y_m),$$

$$\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot$$

$$dy_m/dx = f_m(x, y_1, y_2, \ldots, y_m),$$

in which the functions[1] $f_r(r = 1, 2, \ldots, m)$ are continuous in the neighborhood of $(x_o, y_{1,o}, y_{2,o}, \ldots, y_{m,o})$ which neighborhood is defined by

$$|x - x_o| \leqq a, |y_i - y_{i,o}| \leqq b_i, (i = 1, 2, 3, \ldots, m),$$

let M be the maximum of $|f_1|$, $|f_2|, \ldots, |f_m|$ in the neighborhood defined. Suppose that there exists a set of constants, $K_1, K_2, \ldots, K_m$, such that, for any two points $(x, Y_1, Y_2, \ldots, Y_m)$, $(x, y_1, y_2, \ldots, y_m)$ in the given neighborhood and having the same value for the independent variable x,

$$|f_r(x, Y_1, Y_2, \ldots, Y_m) - f_r(x, y_1, y_2, \ldots, y_m) < K_1|Y_1 - y_1| + K_2|Y_2 - y_2|$$
$$+ \ldots + K_m|Y_m - y_m|,$$ where $r = 1, 2, 3, \ldots, m$. Then, these conditions being satisfied, there exists a unique set of functions

$$y_1 = f_1(x), \; y_2 = f_2(x), \ldots, \; y_m = f_m(x),$$

which, for $|x - x_o| \leqq a'$, a' being the smallest of the $(m+1)$ values a, b_i/M $(i = 1, 2, \ldots, m)$, satisfies the given equations and reduces to

$$y_1 = y_{1,o}, \; y_2 = y_{2,o}, \ldots, \; y_m = y_{m,o} \text{ for } x = x_o.$$ [2]

---

[1]All functions, both the given and the required, considered here and throughout this thesis are single-valued and finite.

[2]Slight changes in this statement would be necessary in the complex domain. (See Chapters III and V).

Even as important as this theorem is in the calculus, the historical facts concerning its development have never been published in English. There is no English translation of the original proofs of the theorem or of the modifications and developments through which they have passed.

Three distinct proofs of the theorem have been developed, two of which are due to Cauchy and one to Picard. The purpose of this thesis is to bring together the historical facts concerning the development and publishing of these proofs and a brief biographical sketch of the great mathematicians who developed them; a translation into the English language of these original demonstrations and the modifications, developments and simplifications by later mathematicians; and, finally, a discussion and comparison of the conditions upon which the various proofs are based and the extent of their generality or applicability.

Following this introduction, one chapter will be devoted to each of the proofs with the final chapter devoted to the general discussion.

CHAPTER II

"THE METHOD OF DIFFERENCE EQUATIONS"

2.1. ORIGINATED BY CAUCHY.--The oldest of the three proofs of the existence theorem for ordinary differential equations has been called "The Method of Difference Equations." It was originated by Augustin Louis Cauchy, who was one of the leaders in insisting on rigorous demonstrations in mathematical analysis.

Cauchy was born at Paris, August 21, 1789, and died at Sceaux, May 23, 1857. In 1805 he entered the École Polytechnique, which was the nursery of so many French mathematicians. Two years later he entered the École des Ponts et Chaussees. From 1810 to 1813 he was engaged as an engineer at Cherbourg. He was a professor at the École Polytechnique from 1816 until 1830, when he went into exile because he was too conscientious to take the oath of allegience demanded of him as a result of the political revolution. In 1838 Cauchy returned to Paris and taught in certain Church schools, the oath demanded of him still preventing his acceptance of a chair in the College de France. When the oath was suspended during the political events of 1848, he again entered the École Polytechnique as Professor of Mathematical Astronomy.

Cauchy was an untiring worker, a man of uncommon scientific ability, a prolific and profound mathematician. From 1830 to 1859 he published more than 600 original memoirs and about 150 reports. In spite of the obscurity, repetition of old results, and blunders caused by his feverish haste, his prompt publication of results and his preparation of standard text-books enabled him to exercise an influence on the great mass of mathematicians that was more immediate and beneficial than that of any contemporary writer. His work includes researches into the theory of residues, the question of convergence, differential equations, the theory of functions, the elucidation of the imaginary, the theory of numbers, operations with determinants, the theory of substitutions, the foundations of calculus, the theory of probability, mathematical astronomy, and the applications of mathematics to physics.

2.2 FACTS AND CIRCUMSTANCES SURROUNDING THE DEVELOPMENT AND PUBLISHING OF THIS PROOF.--It was during his first lectureship at the École Polytechnique in Paris that Cauchy developed, in 1823, the proof we are studying in this chapter. This proof was summarized in a Memoir, "Sur l'integration des Equations Differentielles," lithographed Prague, 1835, and this summary was reprinted in "Exercises d'Analyse," 1840.[1] It was given to the public and preserved for us in its complete form only through the work of Cauchy's friend and pupil, l'abbe Moigno, in his "Lecons de Calculus," 1844. The printing of Moigno's book began in 1841. His original purpose was to publish only one volume, but the abundance of material forced him to alter his plan to include two volumes. The first volume, on the differential calculus, which had already become a fixed science, came out according to schedule. But the second volume, on the integral calculus, was delayed for two reasons. First, Moigno's duties as leader of a monastery demanded part

[1]Ince, "Ordinary Differential Equations" (1927) p. 76 (Footnote).

of his time. Second, and of chief importance, the integral calculus was rapidly changing. A new era seemed to have come. Many noted scientists were working on this branch of mathematics. It required time to analyze and condense the many papers being published by such men as MM. Liouville, Slurin, Binet, Lame, Calalan, Blanchet, and Bertrand in France; MM. Gauss, Jacoby, Lejeune-Dirichlet, and Richelot in Allegmagne; MM. Ostrogradzky and Bouniakowsky in Russia; and M. Tortolini in Italy. At this period Cauchy himself put out more than twenty-four papers about the integral calculus which Moigno wished to analyze in his lessons. Therefore, Moigno's book, which was very modern in its day, was not completed until 1844. Cauchy's proof of the existence theorem by means of difference equations is found in Vol. II, pp. 385-396. In the following section I give a translation of the proof of the theorem and corollaries in their original form. To my knowledge, they have never been published in English.

2·3 A TRANSLATION OF CAUCHY'S ORIGINAL PROOF

2·31. Whenever the differential equation

$$dy = f(x,y)dx$$

is integrable by one of the methods explained in the previous lectures, we can easily obtain, as we have shown, a function of x for the unknown y, which will satisfy the differential equation and will equal y_0 when $x = x_0$. Conversely, the equation

$$dy = f(x, y)dx$$

can be integrated and has a general integral with an arbitrary constant, if one can prove that there is a general value of y which fulfills the two conditions mentioned above. This goal one reaches in most cases by means of the principles about to be set up.

Let X denote a new special value of x and let $x_1, x_2, x_3, \dots, x_{n-1}$ be quantities which lie between the limits x_o and X and which constantly increase or constantly decrease from the first to the second limit. Let us further suppose that by means of the equations

$$y_1 - y_o = (x_1 - x_o) f(x_o, y_o),$$
$$y_2 - y_1 = (x_2 - x_1) f(x_1, y_1),$$
$$\cdots\cdots\cdots\cdots$$
$$Y - y_{n-1} = (X - x_{n-1}) f(x_{n-1}, y_{n-1}),$$

one has calculated n values of y corresponding to $y_o, y_1, \dots, y_{n-1}, Y$, and that one has by eliminating $y_1, y_2, \dots, y_{n-1}$, a value of Y of the form

$$Y = F(x_o, x_1, x_2, \dots, x_{n-1}, X, y_o),$$

which has very remarkable properties. If one now adds all these equations together one has

$$Y - y_o = (x_1 - x_o) f(x_o, y_o) + (x_2 - x_1) f(x_1, y_1) + \dots \quad (X - x_{n-1}) f(x_{n-1}, y_{n-1}).$$

Now, however, the sum in the second part of this last equation equals the product of the sum of the differences $x_1 - x_o$, $x_2 - x_1, \dots, X - x_{n-1}$, or $X - x_o$, and a mean quantity which lies between the coefficients

$$f(x_o, y_o),\ f(x_1, y_1), \dots, f(x_{n-1}, y_{n-1})$$

and if one designates the largest of the absolute values of these coefficients by A, then the mean value will necessarily be expressed by an expression of the form $\pm \Theta A$, where Θ denotes a positive number smaller than unity; one now has

$$Y - y_o = \pm \Theta A (X - x_o)$$
$$Y = y_o \pm \Theta A (X - x_o)$$

and from this it follows that the value of Y must necessarily lie between the limits $y_0 \pm A(X-x_0)$. In like manner we find that the quantities $y_1, y_2, \ldots, y_{n-1}$ lie respectively between the limits

$$y_0 \quad A(x_1 \quad x_0),\ y_0 \quad A(x_2 \quad x_0), \ldots, y_0 \quad A(x_{n-1} \quad x_0),$$

and hence all these quantities, as well as Y, may be reduced to expressions of the form

$$y_0 \pm \Theta A(X \quad x_0).$$

Hence, it follows that the coefficients

$$f(x_0, y_0),\ f(x_1, y_1), \ldots, f(x_{n-1}, y_{n-1})$$

are particular values of the expression

$$f[x_0 + \theta(X-x_0),\ y_0 \pm \Theta A(X-x_0)$$

which correspond to the values of θ and Θ lying between 0 and 1. Now, we will suppose that the largest and the smallest of the coefficients here under discussion correspond respectively to the values of

$$\theta = \theta_0,\ \pm\Theta = \Theta_0;\quad \theta = \theta_0 + e,\ \pm\Theta = \Theta_0 + e',$$

so that every quantity lying between these two coefficients, or between

$$f[x_0 + (\theta_0 + e)(X-x_0),\ y_0 \ [(\Theta_0 + e')A(X-x_0)] \text{ and}$$
$$f[x + \theta_0(X-x_0),\ y_0 + \Theta_0 A(X-x_0)],$$

can be regarded as a particular value of the expression

$$f[x_0 + (\theta_0 + e\zeta)(X-x_0),\ y_0 + (\Theta_0 + e\zeta)A(X-x_0)],$$

which corresponds to a value of ζ lying between the limits 0 and 1, and, hence, as particular value of the expression

$$f[x_0 + \theta(X-x_0),\ y_0 \pm \Theta A(X-x_0)],$$

which corresponds to values of θ and Θ lying between the same limits. It follows that since the difference $Y-y_0$ is equal to the product of $X - x_0$ and a mean quantity of the kind mentioned, one can say that

$$Y - y_0 = (X - x_0) f[x_0 + \theta (X - x_0),\ y_0 \pm \Theta A(X - x_0)]$$

and therefore

$$Y = y_0 + (X - x_0) f[x_0 + \theta (X - x_0),\ y_0 \pm \Theta A(X - x_0)],$$

where θ and Θ again denote two positive numbers less than unity.

2·311. Corollary 1. If all the elements of the difference $X - x_0$, that is, the binomials $x_1 - x_0$, $x_2 - x_1, \ldots, X - x_{n-1}$, were reduced to a single one, which would be this difference itself, then one would have only

$$Y - y_0 = (X - x_0) f(x_0,\ y_0).$$

If one compares this equation with the previous one, then one sees that, through the nature of the division of the interval $X - x_0$ into elements, the second factor of the product, which expresses the value of $Y - y_0$, changes in that the quantities x_0, y_0 in it increase in such a way that their increases are less, respectively, than the numerical value of the first factor and the same when multiplied by the constant A.

2·312. Corollary 2. If m denotes a number less than n and one assumes $X_m = \xi$, $y = \eta$ then one has

$$Y - \eta = (X - \xi) f[\xi + \theta (X - \xi),\ \eta \pm \Theta A (X - \xi)].$$

2·32. After we have learned to know the form of Y, we will also determine in what way this quantity changes with y_0, or calculate the increase $\beta_n \eta Y$ which corresponds to an increase $\beta_0 \eta y_0$. Let $H = +(X - x_0)$ indicate the numerical value of the difference $X - x_0$.

Let us further suppose that when x remains in the limits x_0 and X, the derivative $df(x,y)/dy$ remains continuous with regard to the variables x, y and therefore also lies between the limits C, C being a positive quantity. Now let

$$\mathcal{O}(x,\ y)dx + \chi(x,\ y)dy$$

be the total differential of the function $f(x,y)$ so that one has identically

$$df(x,y)/dx = \mathcal{Q}(x,y),\ df(d,y)/dy = \chi(x,y).$$

Also let $\beta_1, \beta_2, \ldots, \beta_n$ be the respective increases of $y_1, y_2, \ldots, Y$ when one assigns to y_0 the increase β_0, and let $\theta, \Theta_0, \Theta_1, \Theta_2, \Theta_3, \ldots, \Theta_{n-1}$ each be a positive quantity less than unity. Since the equation $y_1 - y_0 = (x_1 - x_0)f(x_0, y_0)$ must hold, if one allows y_0 to increase by β_0 and y_1 by β_1, one then has

$$y_1 + \beta_1 - (y_0 + \beta_0) = (x_1 - x_0)\ f(x_0,\ y_0 + \beta_0),$$

and hence

$$\beta_1 - \beta_0 = (x_1 - x_0)\left[f(x_0,\ y + \beta_0) - f(x_0,\ y_0)\right].$$

Further, one has, by means of a well-known formula in the supposition made,

$$\frac{f(x_0,y_0 + \beta_0) - f(x_0,y_0)}{\beta_0} = \chi(x_0, y_0 \pm \theta\ \beta_0) = \pm\ \Theta_0 C$$

$$f(x_0,y_0 + \beta_0) - f(x_0,y_0) = \pm\ \Theta\ \beta_0 C,$$

and consequently

$$\beta_1 = \beta_0\left[1 \pm \Theta_0 C\ (x_1 - x_0)\right].$$

Similarly one finds

$$\beta_2 = \beta_1 [1 \pm \Theta_1 c(x_2 - x_1)],$$
$$\cdots\cdots\cdots$$
$$\beta_n = \beta_{n-1} [1 \pm \Theta_{n-1} c(X - x_{n-1})],$$
$$\beta_n = \beta_0 [1 \pm \Theta_0 c(x_1 - x_0)][1 \pm \Theta_1 c(x_2 - x_1)] \cdots [1 \pm \Theta_{n-1} c(X - x_{n-1})].$$

If the difference $X - x_0$ is positive, then the numerical value of the binomial $1 \pm \Theta c(x_1 - x_0)$ is less than the sum $1 + c(x_1 - x_0)$ and hence, smaller than the exponential quantity

$$e^{c(x_1 - x_0)} = 1 + c(x_1 - x_0) + \frac{c^2(x_1 - x_0)^2}{1 \cdot 2} + \cdots.$$

For the same reason the numerical values of the binomials

$$1 \pm \Theta_1 c(x_2 - x_1), \ldots, 1 \pm \Theta_{n-1} c(X - x_{n-1})$$

are smaller respectively than the exponential quantities

$$e^{c(x_2 - x_1)}, \ldots, e^{c(X - x_{n-1})};$$

and it follows that the product of all the binomials occurring in the value of β_n is less than the product of all these exponential quantities, that is, less than $e^{c(X - x_0)}$, and is thus reduced to an expression of the form

$$\pm \Theta e^{c(X - x_0)},$$

where again Θ denotes a number lying between 0 and 1. For this expression we would obviously have to substitute

$$e^{c(x_0 - X)},$$

if the difference $X - x_0$ were negative. Accordingly, we have

$$\beta_n = \pm \Theta \beta_0 e^{\pm c(X - x_0)} = \pm \Theta \beta_0 e^{cH},$$

which is the increase β_n of Y, corresponding to the increase β_0 of y_0. If, for the sake of brevity, we substitute K for e^{CH}, K being a positive and finite constant, then we have merely $\beta_n = \pm \Theta K \beta_0$.

2·321. Corollary 1. If the elements $x_1 - x_0$, $x_2 - x_1$, ... of the difference $X - x_0$ all receive numerical values less than $1/C$, then the factors

$$1 \pm \Theta_0 C(x_1 - x_0),\quad 1 \pm \Theta_1 C(x_2 - x_1),\ \ldots$$

are all positive, and we necessarily have

$$\beta_m = \Theta K \beta_0 .$$

2·322. Corollary 2. The value of $\beta_m = \Theta K \beta_0$ becomes infinitely small with β_0; hence, an infinitely small increment of the quantity Y will always correspond to an infinitely small increment of the quantity y_0 and thus the first of these quantities is a continuous function of the second.

2·323. Corollary 3. If one considers only the equations

$$y_{m+1} - y_m = (x_{m+1} - x_m)\ f(x_m, y_m),$$

$$y_{m+2} - y_{m+1} = (x_{m+2} - x_{m+1})\ f(x_{m+1}, y_{m+1})$$

$$\cdots\cdots\cdots\cdots\cdots\cdots$$

$$Y - y_{n-1} = (X - x_{n-1})\ f(x_{n-1}, y_{n-1}),$$

they are sufficient for the determination of Y as a function of the quantities x_m, x_{m+1}, ..., x_{n-1}, X, y_m, and one can again easily show that, if one assigns to y_m a certain increment, the corresponding increment of Y is of the following form

$$\pm \Theta \beta_m e^{\pm C(X - x_m)}.$$

Hence, this last increment has a lower numerical value than that of the product $\beta_m e^{\pm C(X - x_m)}$ and, even more so, than that of the product

$$\beta_m e^{C(X - x_0)} = K\beta_m .$$

2.33. The quantity Y is obviously dependent, (1) on the limiting values x_0, X; (2) on the quantity y_0; and (3) on the number n and the values of the elements into which the difference $X - x_0$ is divided, or, in other words, on the chosen manner of dividing the difference. It can be shown, however, that the value of Y is dependent merely on the three quantities x_0, X and y_0, if one lets the numerical values of the elements of the difference $X - x_0$ approach the limit 0 by increasing their number indefinitely. To this end, one needs only prove that the chosen manner of division no longer has any perceptible influence on the value of Y if the number n becomes very large, which one can easily do in the following manner.

If the elements of the difference $X - x_0$ reduce themselves to a single one, which then becomes the difference itself, then the value of Y is determined by the equation

$$Y - y_0 = (X - x_0)\ f(x_0,\ y_0).$$

If, on the other hand, this difference $X - x_0$ is divided into n elements

$$x_1 - x_0,\ x_2 - x_1, \ldots, X - x_{n-1},$$

we have then

$$Y - y_0 = (X - x_0)\ f\left[x_0 + \theta\ (X - x_0),\ \ y_0 \pm \Theta\ \Lambda\ (X - x_0)\right].$$

In order to proceed to a second manner of division we need only to divide each element of the first division into new elements, and one can then by approximation calculate the influence which each subdivision has on the value of Y. For, if one, for example, divides the element $x_1 - x_0$ into several parts, then we have for the equation

$$y_1 - y_0 = (x_1 - x_0)\ f(x_0,\ y_0)$$

several other equations of the same form; but if one proceeds in the well-known manner, one finds

$$y_1 - y_0 = (x_1 - x_0)\, f[x_0 + \theta (x_1 - x_0),\, y_0 \pm \Theta A]\, (x_1 - x_0)]),$$

where θ and Θ again denote two positive numbers less than unity.

If one supposes

$$f[x_0 + \theta (x_1 - x_0),\, y_0 \pm \Theta A (x_1 - x_0)] = f(x_0, y_0) \pm \epsilon_0,$$

one has

$$y_1 - y_0 = (x_1 - x_0)\, f(x_0, y_0) \pm \epsilon_0 (x_1 - x_0).$$

But before the further division of the element $x_1 - x_0$ one had

$$y_1 - y_0 = (x_1 - x_0)\, f(x_0, y_0),$$

and, hence, the value of y_1, by this new division, is changed by the product $\pm \epsilon_0 (x_1 - x_0)$.

If, however, the remaining elements of the difference $X - x_0$ retain their original values while the quantity y_1 receives the increment $\pm \epsilon_0 (x_1 - x_0)$; then Y receives, according to the statement above, another increment of the form

$$\pm \Theta K \epsilon_0 (x_1 - x_0).$$

Hence, the increase of Y caused by the subdivision of the single element $x_1 - x_0$ has a numerical value less than the quantity

Similarly, one proves that the increase of Y caused by the subdivision of the element $x_{m+1} - x_m$ has a lesser numerical value than the quantity

$$\pm K \epsilon_m (x_{m+1} - x_m),$$

where the number ϵ_m is determined by an equation of the form

$$\pm \epsilon_m = f[x_m + \theta (x_{m+1} - x_m), y_m \pm \Theta A (x_{m+1} - x_m)] - f(x_m, y_m).$$

If one thus successively divides anew all elements of the difference $X-x_o$, then Y receives a series of increments, whose sum is less than

$$K\epsilon_o(x_1-x_o)+K\epsilon_1(x_2-x_1)+\cdots+K\epsilon_{n-1}(X-x_{n-1})=K\epsilon(X-x_o),$$

where ϵ denotes a middle quantity among the numbers $\epsilon_o, \epsilon_1, \epsilon_2, \ldots$. If the differences x_1-x_o, x_2-x_1,...become infinitely small, then the same is true of the quantities $\epsilon_o, \epsilon_1, \epsilon_2, \ldots$ as well as of the expression

$$K\epsilon(X-x_o);$$

and, therefore, the value of Y, corresponding to a certain division in which the elements of the difference $X-x_o$ have very small numerical values, will not be perceptibly changed, if one proceeds to a second division in which each of the elements is again divided into several others.

Let us now assume that one observes at the same time two divisions of such a nature that the elements of the second division are no longer subdivisions of the elements of the first division. Then one can compare both these divisions with a third of such a nature that each element of the first or second division is formed by the union of several elements of the third. For the fulfillment of this condition nothing further is required except that all values of x inserted in the first divisions are also used in the third, and one can prove that one changes the value of Y very little when one proceeds from the first or second division to the third. If, therefore, the elements of the difference $X-x_o$ become infinitely small, then the manner of subdivision no longer has any perceptible influence on the value of Y, and if one lets the numerical values of these elements decrease, by increasing their number infinitely, then the value of

Y converges toward a certain limit which depends only on the form of the function $f(x,y)$, on the limit values X and x_0 of the variable x and on the quantity y_0.

2·331. Corollary 1. <u>Since the limit toward which Y converges, when the elements of the difference</u> $X-x_0$ <u>become infinitely small, depends only on the three values</u> x_0, X <u>and</u> y_0, <u>let us denote this limit by</u> $F(x_0, X, y_0)$; and by $F(X, y_0)$, $F(X)$, <u>if we wish to allow only the two quantities</u> X, y_0, <u>or the one quantity</u> X <u>to be changed.</u>

2·34. It can now be easily proved that there always exists one function of x, which satisfies the differential equation $dy = f(x,y)dx$ and which assumes a special but arbitrary value y_0 if one assigns to the variable x a given value x_0. For let $F(x)$ be the value of $F(X)$ if one substitutes x for X. Since $F(X)$ is the value of Y, if the elements of the difference $X-x_0$ become infinitely small, then we have from, the equation

$$Y-\eta = (X-\xi)\, f[\xi+\theta(X-\xi),\ \eta \pm \Theta A(X-\xi)],$$

the equation

$$F(X)-F(\xi) = (X-\xi)\, f[\xi+\theta(X-\xi),\ F(\xi)\pm\Theta A(X-\xi)];$$

and if one substitutes

$$\xi = x,\quad X = x+h$$

where x and $x+h$ lie between the limits x_0 and X, then we have at the same time

$$F(x) = y_0 + (x-x_0)\ f[x_0+\theta(x-x_0),\ \ y_0 \pm \Theta A(x-x_0)]$$

and

$$F(x+h)-F(x) = hf[x+\theta h,\ \ F(x)\pm\Theta Ah].$$

Now it is easy to see that (1) when $x = x_0$, $F(x)$ reduces to y_0; (2) if h becomes infinitely small, the corresponding increase of the function $F(x)$, namely, $F(x+h) - F(x)$, also is an infinitely small quantity; and (3) that from this equation divided by h the following results:

$$F'(x) = f\left[x,\ F(x)\right];$$

which expresses the fact that the function $F(x)$ satisfies the differential equation $dy = f(x,y)dx$.

Therefore, if the function $f(x,y)$ and its derivative $\frac{df(x,y)}{dy}$ remain finite and continuous between the limits x_0, and X, then there exists a function of x which satisfies the differential equation $dy = f(x,y)dx$, and which assumes the value y_0, if one assigns to the variable x a given value x_0.

2.341. Corollary 1. If one designates the limit of Y by $F(x,y_0)$ then the function y appears in the form of $y = F(x,y_0)$ and is the general integral of the given differential equation, because y_0 is an arbitrary constant, and this integral also, like Y, is a continuous function of y_0.

2·4. IMPROVEMENT DUE TO LIPSCHITZ.--In a paper published in the "Bulletin des Sciences Mathematique et Astronomiques" (1) 10 (1876), pp. 149-159, M. R. Lipschitz has greatly improved Cauchy's proof by making prominent the conditions upon which it is based. This improvement by Lipschitz has become almost as famous as the original proof by Cauchy. Following is a translation of the improvement due to Lipschitz taken from the paper mentioned above.

We shall suppose a given system of differential equations, with x as the independent variable and $y^1, y^2, \ldots, y^n$ as the dependent variables, in the form

(1) $$dy^\alpha/dx = f^\alpha(x, y^1, y^2, \ldots, y^n),$$

where $\alpha = 1, 2, \ldots, n$. The functions f^α are given for connected set of values of the variables $x, y^1, y^2, \ldots, y^n$. [1]This set of values is said to be the domain G. (If $n = 2$, we can consider x, y^1, y^2 as representing a point in space, and we have a very nice image of G.) For all points in the domain G the n functions f^α are to be uniform, continuous and bounded. Moreover, they must satisfy the following inequality

(2) $$\begin{cases} |f^\alpha(h, k^1, k^2, \ldots, k^n) - f^\alpha(h, l^1, l^2, \ldots, l^n)| \\ < C^{\alpha,1}|k^1 - l^1| + C^{\alpha,2}|k^2 - l^2| + \cdots + C^{\alpha,n}|k^n - l^n|, \end{cases}$$

for any two points $x = h, y^\alpha = k^\alpha$ and $x = h, y^\alpha = l^\alpha$, having the same value for the independent variable. The quantities $C^{\alpha,\beta}$ are positive constants and here, as in the following, the symbol $|w|$ represents the absolute value

[1] A domain G is said to be connected when it is possible to join any two points whatever of that domain by a continuous path which lies entirely in that region of the plane. (Goursat-Hedrick-Dunkel, "Functions of A Complex Variable", Vol.II, Part I, p.11).

of ω. The imposed condition of continuity demands that for any two systems of values $x = h$, $y^\alpha = k^\alpha$ and $x = j$, $y^\alpha = l^\alpha$, the difference

$$|f^\alpha(h, k^1, k^2, \ldots, k^n) - f^\alpha(j, l^1, l^2, \ldots, l^n)|$$

can be made as small as we please if the differences $|h - j|$, $|k^\alpha - l^\alpha|$ approach zero. If we consider inequality (2) we see that this condition assumes that we can choose the difference $|h - j|$ so small that it has, as a consequence, the inequality

(3) $$|f^\alpha(h, l^1, l^2, \ldots, l^n) - f^\alpha(j, l^1, l^2, \ldots, l^n)| < \sigma$$

howeoever small σ may be.

The system (1) will be completely integrated if we determine a system of functions y^1, $y^2, \ldots, y^n$ satisfying equations (1) and for a given value of x, $x = x_0$, satisfying the equations

(1) $$y^\alpha = y^\alpha_0$$

The system of values $(x_0, y^1_0, y^2_0, \ldots, y^n_0)$ must be included in the domain G, exclusive of the boundary. So that we may find positive quantities a_0, b^α_0, such that the inequalities $|x - x_0| \leqq a_0$, $|y^\alpha - y^\alpha_0| \leqq b^\alpha_0$ define a subdomain of G; consequently, there exist positive finite constants c^α_0 such that, for all points of the subdomain, we shall have

(4) $$|f^\alpha| < c^\alpha.$$

If we determine the positive quantity A_0 so that we have

(4) $$A_0 c^\alpha_0 < b^\alpha_0, \qquad A_0 < a_0,$$

the domain described by the inequalities

(4) $$|x - x_0| \leqq A_0, \quad |y^\alpha - y^\alpha_0| \leqq b^\alpha_0$$

will lie entirely within G. We shall call it H_0.

<u>Under these conditions, there always exists a unique system of functions $y^1, y^2, \ldots, y^n$, satisfying equations (1) and varying continuously within H_o if the variable x passes from $x_o - A_o$ to $x_o + A_o$, and satisfying the equations $y = y_o$ when $x = x_o$.</u>

For the proof of this theorem it is sufficient to consider only those values of x in the interval $x_o \leq x \leq x_o + A$ for the proof in the interval $x_o \geqslant x \geqslant x_o + A$ can be developed in the same way. Suppose that there is between x_o and $x_o + A$ a sequence of intermediary values $x_1, x_2, \ldots, x_{p-1}$, such that

$$x_o < x_1 < x_2 < \cdots < x_p = x_o + A$$

and determine n quantities $y\eta_1$, by the n equations

$$(5) \qquad \eta_1^\alpha - y_o^\alpha = f^\alpha(x_o, y_o^1, y_o^2, \ldots, y_o^n)\,(x_1 - x_o).$$

These equations would be the same as the given differential equations (1) if we replace, in the left hand member of them, dx and dy by the finite quantities $x_1 - x_o$, $\eta_1^\alpha - y_o^\alpha$ respectively, and in the right hand member $x_1 y^\alpha$ by x_o, y_o^α. From the inequalities (4) and (4^α) we can conclude that the equations (5) have as consequence the inequalities

$$\left|\eta_1^\alpha - y_o^\alpha\right| < c_o^\alpha (x_1 - x_o) < b_o^\alpha,$$

and, therefore, the system of values $(x_1, \eta_1^1, \eta_1^2, \ldots, \eta_1^n)$ is in the domain H_o. In the same manner we can form the sequence of systems of values

$$(x_{a+1}, \eta_{a+1}^1, \ldots, \eta_{a+1}^n)$$

by substituting successively $a = 1, 2, \ldots, p-1$ in the equation

$$(5^\alpha) \qquad \eta_{a+1}^\alpha - \eta_a^\alpha = f^\alpha(x_a, \eta_a^1, \eta_a^2, \ldots, \eta_a^n)(x_{a+1} - x_a).$$

All the systems of values will certainly remain in the domain H_o. We continue the decomposition of the interval by interpolating between x_a and x_{a+1} the $q_a - 1$ values of x,

$$x_{a,1}, x_{a,2}, \dots, x_{a,q_a-1} < x_{a,q_a} = x_{a+1,0} = x_{a+1}.$$

By this new division we obtain a new sequence of systems of values beginning with $(x_0, y_0^1, y_0^2, \dots, y_0^n)$ each of which will be included in H_0. This new sequence of values $(x_{a,\mu_a}, \eta'_{a,\mu_a}, \dots \eta^n_{a,\mu_a})$ will be obtained by replacing in the equation

(6) $$\eta^\alpha_{a,\mu_a+1} - \eta^\alpha_{a,\mu_a} = f^\alpha(x_{a,\mu_a}, \eta'_{a,\mu_a}, \eta^2_{a,\mu_a}, \dots, \eta^n_{a,\mu_a})(x_{a,\mu_a+1} - x_{a,\mu_a}),$$

a by 0, 1, 2,...,p−1, μ_a by the numbers 0, 1, 2,..., q_a-1, and, finally, by placing

$$\eta^\alpha_{0,0} = y^\alpha_0, \quad \eta^\alpha_{a,q_a} = \eta^\alpha_{a+1,0}.$$

It is now our task to establish that, regarding the first quantities $x_1, x_2, \dots, x_{p-1}$ as fixed, increasing the number q_a indefinitely, and indefinitely decreasing the intervals by any law whatsoever, the values

$$\eta^\alpha_{a+1,0} = y^\alpha_{a+1},$$

corresponding to the value $x = x_{a+1}$ of the independent variable, converge to a fixed limit independent of the law of increasing q_a and the law of decreasing the secondary intervals into which we have divided the intervals between the quantities $x_0, x_1, x_2, \dots, x_{p-1}, x_p$. This proof shows that under the conditions given above it is always possible to choose the quantities $x_1, x_2, \dots, x_{p-1}$ so that the absolute value of the differences $y^\alpha_{a+1} - \eta^\alpha_{a+1}$ remain, for each a, smaller than a quantity σ, chosen arbitrarily small.

If in equations (6) we set successively $\mu_a = 0, 1, 2, \dots, \mu'_a$ and add them, we obtain the following

(7) $$\eta^\alpha_{a,\mu_a+1} - \eta^\alpha_{a,\mu_a} = \sum_{\mu'_a=0}^{\mu'_a=\mu_a} f^\alpha(x_{a,\mu'_a}, \eta'_{a,\mu'_a}, \dots, \eta^n_{a,\mu'_a})(x_{a,\mu'+1} - x_{a,\mu'_a}),$$

which gives by means of inequality (4),

$$|\eta^{\alpha}_{a,\mu_a+1}-\eta^{\alpha}_{a,0}|<C^{\alpha}_{0}(x_{a,\mu_a+1}-x_{a,0})<C^{\alpha}_{0}(x_{a+1}-x_{a}).$$

This inequality expresses that the system of values ($x_{a,\mu_a}, \eta^1_{a,\mu_a}, \eta^2_{a,\mu_a}, \ldots, \eta^n_{a,\mu_a}$) remains, if we let a be fixed and if we give μ_a all the values from o to $q_a - 1$, in the domain K_0 whose limits with respect to each of the $n+1$ variables can be made as close together as we please by taking the difference $x_{a+1} - x_a$ sufficiently small. As, by hypothesis, the functions f^{α} remain continuous in the domain H_0, the difference $x_{a+1} - x_a$ can be chosen so small that in the domain K_0 the difference between two values of this function will become smaller than a quantity λ, chosen arbitrarily small. Supposing that the difference is determined in this way, by taking the number p itself sufficiently large, we obtain from equation (7) when $\mu_a = q_a - 1$, the result

$$(8) \quad y^{\alpha}_{a+1} - y^{\alpha}_{a} = |f^{\alpha}(x_a, y^1_a, y^2_a, \ldots, y^n_a) + \varepsilon^{\alpha}_a \lambda|(x_{a+1} - x_a),$$

where the quantities ε^{α}_a are proper fractions, either positive or negative.

Subtracting the equation (5) from this equation we get

$$(9) \quad y^{\alpha}_{a+1} - \eta^{\alpha}_{a+1} = y^{\alpha}_a - \eta^{\alpha}_a + |f^{\alpha}(x_a, y^1_a, \ldots, y^n_a) - f^{\alpha}(x_a, \eta^1_a, \ldots \eta^n_a) + \varepsilon^{\alpha}_a \lambda|(x_{a+1} - x_a);$$

but, because of (2), we shall have

$$|f^{\alpha}(x_a, y^1_a, \ldots, y^n_a) - f^{\alpha}(x_a, \eta^1_a, \ldots, \eta^n_a)| < C^{\alpha,1}|y^1_a - \eta^1_a| + C^{\alpha,2}|y^2_a - \eta^2_a| + \cdots + C^{\alpha,n}|y^n_a - \eta^n_a|.$$

If one then places

$$(10) \quad |y^{\alpha}_a - \eta^{\alpha}_a| = Z^{\alpha}_a$$

the equations (9) give the sequence of inequalities

(11) $$Z^{\alpha}_{a+1} < Z^{\alpha}_{a} + (c^{\alpha,1} Z^{1}_{a} + c^{\alpha,2} Z^{2}_{a} + \cdots + c^{\alpha,n} Z^{n}_{a} + \lambda)(x_{a+1} - x_{a}).$$

Moreover we have

(11^{α}) $$Z^{\alpha}_{0} = 0$$

Now it is clear that if we form a sequence of quantities by means of the equations

(12) $$\begin{cases} u^{\alpha}_{a+1} - u^{\alpha}_{a} = (c^{\alpha,1} u^{1}_{a} + c^{\alpha,2} u^{2}_{a} + \cdots + c^{\alpha,n} u^{n}_{a} + \lambda)(x_{a+1} - x_{a}) \\ u^{\alpha}_{a} = 0 \end{cases}$$

in which the index a takes values from 0 to p— 1, we will constantly have

(13) $$Z^{\alpha}_{a+1} < u^{\alpha}_{a+1}.$$

Now we need, for our demonstration, to show that, taking λ sufficiently small, the quantities Z^{α}_{a+1} remain as small as one would desire; our purpose will be attained if we prove the same thing for the quantities u^{α}_{a+1} or for greater quantities.

Let c be a positive quantity superior to the greatest of the n^2 constants $c^{\alpha,\beta}$; if one determines the quantities V^{α}_{a} by the equations

(12^{α}) $$\begin{cases} V^{\alpha}_{a+1} - V^{\alpha}_{a} = |c(V^{1}_{a} + V^{2}_{a} + \cdots + V^{n}_{a}) + \lambda|(x_{a+1} - x_{a}), \\ V^{\alpha}_{0} = 0, \end{cases}$$

one will evidently have, for $\alpha = 1, 2, 3, \ldots, p-1$,

(13^{α}) $$u^{\alpha}_{a+1} < V^{\alpha}_{a+1}.$$

Now the first equation (12^{α}) gives

$$V^{1}_{a+1} - V^{1}_{a} = V^{2}_{a+1} - V^{2}_{a} = \cdots = V^{n}_{a+1} - V^{n}_{a},$$

and, by virtue of the second, we have

$$V^{1}_{a} = V^{2}_{a} = V^{3}_{a} = \cdots = V^{n}_{a}.$$

The first equation (12^{α}) may then be written

$$V^{\alpha}_{a+1} - V^{\alpha}_{a} = (ncV^{\alpha}_{a} + \lambda)(x_{a+1} - x_{a}),$$

or

$$V^{\alpha}_{a+1} + \lambda/nc = \left(V^{\alpha}_{a} + \frac{\lambda}{nc}\right)\left|1 + nc(x_{a+1} - x_a)\right|;$$

one then has

$$(14) \qquad V^{\alpha}_{a+1} = -\frac{\lambda}{nc} + \frac{\lambda}{nc}\left|1+nc(x_1 - x_0)\right|\left|1+nc(x_2 - x_1)\right|\cdots\left|1+nc(x_{a+1} - x_0)\right|.$$

But the product

$$\left|1+nc(x_1 - x_0)\right|\left|1+nc(x_2 - x_1)\right|\cdots\left|1+nc(x_{a+1} - x_a)\right|,$$

in which nc is positive, has itself a positive value inferior to

$$e^{nc(x_{a+1} - x_0)}$$

we then have

$$V^{\alpha}_{a} < -\frac{\lambda}{nc} + \frac{\lambda}{nc}e^{nc(x_{a+1} - x_0)}.$$

The comparison of this inequality with the inequalities (13) and (13a) gives

$$(15) \quad \left|y^{\alpha}_{a+1} - \eta^{\alpha}_{a+1}\right| = Z^{\alpha}_{a+1} < \frac{-1+e^{nc(x_{a+1}-x_0)}}{nc}\lambda < \frac{-1+e^{ncA_0}}{nc}\lambda,$$

and since the factor

$$\frac{-1+e^{ncA_0}}{nc}$$

has a finite value, one concludes from it that the difference

$$\left|y^{\alpha}_{a+1} - \eta^{\alpha}_{a+1}\right|$$

may be made as small as one would desire; because λ is a quantity as small as one desires, depending only on the choice of the intervals $x_1 - x_0$, $x_2 - x_1 \ldots, x_p - x_{p-1}$. Since the differences $\left|y^{\alpha}_{a+1} - \eta^{\alpha}_{a+1}\right|$ may be taken as small as one desires, the quantities y^{α}_{a+1} which correspond to the fixed value $x = x_{a-1}$ of the variable, converge toward a determined limit, independent of the law of increase of the number q_a, and of the law of decrease of the new intervals. These limited values, by virtue of equations (8), define a system of solutions of the differential equations, a system for which the functions y^α are reduced to y_0^α when $x = x_0$. The existence of a system of solutions satisfying the imposed conditions is then established, and the first part of our program is fulfilled.

That there exists no other solution of the system (1) satisfying the stated conditions, one sees as follows: let $y^\alpha = Y^\alpha$ be such a solution, the interval of x_0, $x_0 - A$, to which one may still be limited, may be divided, by the introduction of the quantities $x_1\ x_2, \ldots, x_{p-1}$ in intervals for which, ε_a^α being a proper fraction, positive or negative, and λ a quantity as small as one would wish, one will have the equation

(16) $$\left(Y_{a+1}^\alpha - Y_a^\alpha\right) = \left| f^\alpha(x_a, Y_a', \ldots Y_a^n) + \varepsilon_a^\alpha \lambda \right| (x_{a+1} - x_a),$$

in which the values $y^\alpha = Y^\alpha$ correspond to the value $x = x_a$ of the independent variable, and in which one has, by hypothesis, $Y_0^\alpha = y_0^\alpha$. This sequence of equations is an immediate consequence of the two hypotheses by virtue of which the given system of functions $y^\alpha = Y^\alpha$ satisfies the differential equations (1), and varies continuously when the variable x goes from x_0 to $x_0 + A$. If, by the equations (5) and (5^α), one forms the quantities η_a^α according to the values $x_0, x_1, \ldots, x_{p-1}$, one recognizes that the differences

$$\left| Y_{a+1}^\alpha - \eta_{a+1}^\alpha \right|$$

should behave like the differences

$$\left| y_{a+1}^\alpha - \eta_{a+1}^\alpha \right|,$$

because the equation (16) is deducted from the equation (8) by replacing y_a^α by Y_a^α and y_{a+1}^α by Y_{a+1}^α, and one has besides $Y_0 = y_0$: it results from this that the differences $\left| Y_{a+1}^\alpha - \eta_{a+1}^\alpha \right|$ may be made, like the differences $\left| y_{a+1}^\alpha - \eta_{a+1}^\alpha \right|$, as small as one desires; but, because of the inequality

$$|Y^\alpha_{a+1} - y^\alpha_{a+1}| \leq |Y^\alpha_{a+1} - \eta^\alpha_{a+1}| + |y^\alpha_{a+1} - \eta^\alpha_{a+1}|,$$

the difference $|Y^\alpha_{a+1} - y^\alpha_{a+1}|$ may evidently be made as small as one desires; as a consequence, the given system Y^α_a of the functions Y^α_a may not differ from the system of the functions y^α_a obtained by the division of the first intervals into new intervals that one has made decrease indefinitely. Our demonstration is thus entirely completed.

There is reason to make some remarks on the conditions imposed on the functions f^α. We have supposed in giving a sufficiently small increase to the independent variable and in keeping the same values for the y, one could satisfy the conditions of continuity

(3) $$|f^\alpha(h, l', l'', \ldots, l^n) - f^\alpha(j, l', l'', \ldots, l^n)| < \sigma,$$

σ being taken as small as one desires; moreover, we have imposed, for variations of the quantities y, the conditions of a special nature defined by the inequalities

(2) $$\begin{cases} |f^\alpha(h, k', k'', \ldots, k^n) - f^\alpha(h, l', l'', \ldots, l^n)| \\ < C^{\alpha,1}|k' - l'| + C^{\alpha,2}|k'' - l''| + \cdots + C^{\alpha,n}|k^n - l^n|. \end{cases}$$

<u>Our demonstration supposes essentially that these conditions are satisfied.</u> In order to be convinced of it, one may limit oneself to the case of a single differential equation

$$dy/dx = f(x, y).$$

If, instead of the inequality (2), one supposes only that one has

(2*) $$|f(h, k) - f(h, l)| < C|k - l|^\delta$$

in which δ is a positive quantity smaller than 1, the reasonings, copied closely from those which we have made, lead to replacing the inequality (11) by the following:

(11*) $$Z_{a+1} < Z_a + (cZ_a^\delta + \lambda)(x_{a+1} - x_a)$$

still taking $Z_0 = 0$.

In order to learn whether the quantities Z_{a+1} may, for a value of λ sufficiently small, be made as small as one desires, one will form the system of equations

(12*) $$u_{a+1} - u_a - (cu_a^\delta + \lambda)(x_{a+1} - x_a) = 0,$$

and one will have then

(13*) $$Z_{a+1} < u_{a+1};$$

one must find out then whether or not, for a value of λ sufficiently small, the quantities u_a may be made as small as one would desire. In the first case, one would arrive at the same conclusion; but, in the second case (and we are going to see that it is with this one that we must deal), our demonstration collapses. One has supposed the intervals $x_1 - x_0$, $x_2 - x_1, \ldots, x_p - x_{p-1}$ small enough so that, in a domain K_0, in which the difference of the values of x is less than $x_{a+1} - x_a$, the difference of the two values of y is less than the given quantity λ.

The intervals $x_1 - x_0, x_2 - x_1, \ldots, x_p - x_{p-1}$ being chosen arbitrarily, nothing prevents subdividing them into smaller intervals; the assumed condition remains fulfilled. By carrying the subdivision far enough and designating by u_a the value corresponding to x_a, one sees that the equations (12*) may be replaced, with an approximation as great as one desires, by the equation

$$x_a - x_0 = \int_0^{u_a} \frac{du}{cu^\delta + \lambda},$$

from which

$$x_a - x_0 < \int_x^{u_a} \frac{du}{cu^\delta} = \frac{1}{c(+1-\delta)} u_a^{-\delta+1},$$
$$u_a > (-\delta+1)c(x_a - x_0)^{\overline{-\delta+1}}.$$

Then since λ is a fixed quantity, one can by no means make u_a as small as one desires; the inequalities (13*) cannot then lead the quantities Z_{a+1} to a proper limit.

It is clear that the inequalities (2) are always satisfied when the functions f^α, for all the points of the domain G, have partial derivatives of the first order, uniform, finite, and continuous in relation to the n variables y^α; because then the difference

$$f^\alpha(h, k_1', k_1^2 \ldots k^n) - f^\alpha(h, l_1', l_1^2, \ldots, l^n)$$

may, in accordance with Taylor's theorem, be placed under a formula which makes these inequalities evident. On the other hand, one can conclude nothing on the nature of partial derivatives from these supposedly true inequalities.

In the case where the functions f^α do not contain the variables y^α, the functions remain uniform, finite, and continuous in relation to x; our analysis shows that the integral $\int_0^x f^\alpha(\xi)d\xi$ has a determined sense and that the derivative of this function, taken for a value of the variable equal to the upper limit of the integral, is equal to f (x). The posthumous memoirs of Riemann, on the representation of a function by a trigonometrical series, has thrown light upon this fact, that the existence of the definite integral depends upon a more general condition than continuity: the integral $\int_{x_0}^{x_0+A} f^\alpha(\xi)\,d\xi$ will exist if the function $f^\alpha(x)$ remains finite when x varies from x_0 to x_0+A, and if by dividing the interval from x_0 to x_0+A into intervals indefinitely decreasing,

x_1 x_0, x_2 x_1,..., the total sum of the intervals for which the oscillation of the function $f^{\alpha}(x)$ remains less than a given quantity σ, as small as one would desire, can be made as small as one would desire if these conditions are fulfilled, and if x is a quantity between x_0 and $x_0 + A$, it is clear that the integral $\int_{x_0}^{x} f^{\alpha}(\xi)d\xi$ will exist; but, as it appears to me, these conditions entail in no way this consequence, that the derivative of this integral, is equal to $f^{\alpha}(x)$: and so I thought I should retain the condition of continuity of the functions $f^{\alpha}(x)$ for the study of the integration of differential equations.

Note: In the final chapter of this thesis there are listed the names and works of other mathematicians who have made noteworthy investigations of the method of difference equations.

CHAPTER III

"THE CALCULUS OF LIMITS"

3.1. ORIGINATED BY CAUCHY.--Cauchy's first proof, which was the subject of the preceding chapter, is for real functions and real variables. He has given a second proof, which he called "The Calculus of Limits," for complex variables. [1]This proof was published in the lithographed memoirs of Turin (October, 1831, 1832, and March, 1833), and of Prague (1835). The first of these memoirs was reproduced, in part, in the "Exercises d'Analyse et de Phys. Math." 2, Paris, 1841, p. 41; the second was reproduced in the "Exercises d'Analyse et de Phys. Math." 1, Paris, 1840, p. 327. Extensive notes on this theorem were later published by Cauchy in the "Comptes Rendus Academie of Sciences, Paris," 9-11, 14, 15, 23 (1839-46) and many of these notes were republished in "Cauchy's Collected Works," Series 1, Volumes 4-7 and 10. The most important of them bear the following dates: November 5 and 21, 1839; June 29, October 26, November 2 and 9, 1840; June and July, 1842; September, October, 1846.[2] In the following section is a statement of Cauchy's original demonstration by means of the calculus of limits translated from Vol. 7 of his "Collected Works."

---

[1]P. Painlevé, "Encyclopedie des Sciences Mathematiques" t. II, Vol.3, p. 16.

[2]See also "C. R. Acad. sc. Paris" 9 (1839), p.184-90, 637-49; 10 (1840), p. 957-69; 11 (1840), p. 639-58, 667-77; 730-46; 14 (1842), p. 1020-8; 15 (1842), p. 14-25; 23 (1846), p. 485-7; 529-37, 563-9, 617-9, 702-4, 729-40, 779-87; "Collected Works" (1) 4, Paris 1884, p. 483-91; (1) 5, Paris 1885, p. 5-20, 236-49, 360-409; (1) 6, Paris 1888, p. 461-7; (1) 7, Paris 1892, p. 5-17; (1) 10, Paris 1897, p. 107-9, 124-33, 143-50, 150-3, 169-71, 171-86, 186-96.

3.2. A TRANSLATION OF CAUCHY'S ORIGINAL PROOF.

3.21. General Considerations.

Let us be given a system of differential equations of the form

(1) $D_t x = X, \quad D_t y = Y, \quad D_t z = Z, \ \ldots \ ,$

with the independent variable t and the unknowns x, y, z,... . Let X, Y, Z,... designate given functions of the unknowns x, y, z,...,t. On the other side, let ξ, η, ζ ,... be the new values which the unknowns x, y, z,..., acquire when the variable t acquires a new value designated by τ. If one substitutes

$$\xi \text{ for } x, \quad \eta \text{ for } y, \quad \zeta \text{ for } z, \ldots,$$

a given function

(2) $R = F(x,y,z,\ldots)$

of the unknowns x, y, z,... will have a new value represented by

$$F(\xi, \eta, \zeta, \ldots);$$

and if this new value can be expanded according to Taylor's formula into a convergent series of ascending powers of the difference $(\tau - t)$, then one will have

(3) $F(\xi, \eta, \zeta, \ldots) = R + \frac{\tau - t}{1} D_t R + \frac{(\tau - t)^2}{1.2} D_t^2 R + \cdots\cdots .$

If in equation (2) one replaces successively the function F(x,y,z,...) by each of the unknowns x, y, z,..., one shall have the formulas

(4)
$$\begin{aligned}
\xi &= x + \frac{\tau - t}{1} D_t x + \frac{(\tau - t)^2}{1.2} D_t^2 x + \cdots, \\
\eta &= y + \frac{\tau - t}{1} D_t y + \frac{(\tau - t)^2}{1.2} D_t^2 y + \cdots, \\
\zeta &= z + \frac{\tau - t}{1} D_t z + \frac{(\tau - t)^2}{1.2} D_t^2 z + \cdots, \\
&\cdots\cdots\cdots\cdots\cdots,
\end{aligned}$$

which will represent the solutions of equations (1) whenever the sequences

$$
(5) \qquad \begin{array}{l}
x, \ \frac{(\tau - t)}{1} D_t x, \ \frac{(\tau - t)^2}{1 \cdot 2} D_t^2 x \cdots, \\
y, \ \frac{(\tau - t)}{1} D_t y, \ \frac{(\tau - t)^2}{1 \cdot 2} D_t^2 y \cdots, \\
. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad ,
\end{array}
$$

are convergent. That, at least, one can easily demonstrate by the aid of a general theorem which I have found about the expansion of functions into series.[1] Therefore, to establish the existence of general integrals of equations (1), it will be sufficient to prove that one can assign to $\tau - t$ an absolute value sufficiently small to make the sequences (5) convergent, all of which are special cases of the more general sequences

$$
(6) \qquad R, \ \frac{\tau - t}{1} D_t R, \ \frac{(\tau - t)^2}{1 \cdot 2} D_t^2 R, \ . \ . \ . \ .
$$

Therefore, if one designates the absolute value of $\tau - t$ by ι and by

$$
\mathfrak{I}_0, \ \mathfrak{I}_1, \ \mathfrak{I}_2, \ . \ . \ .
$$

the upper limits of the absolute values of the quantities

$$
R, \ \frac{1}{1} D_t R, \ \frac{1}{1 \cdot 2} D_t^2 R, \ . \ . \ . ,
$$

it will be sufficient to prove that the absolute value ι can be made so small as to render the sequence

$$
(7) \qquad \mathfrak{I}_0, \ \mathfrak{I}_1 \iota, \ \mathfrak{I}_2 \iota, \ . \ . \ . \ .
$$

convergent.

Now let us note that, from formula (2) connected with equations (1), we shall have

$$
D_t R = D_x F(x, y, z, \ldots) X + D_y F(x, y, z \ldots) Y + \cdots
$$

---

[1] Goursat-Hedrick-Dunkel, "Mathematical Analysis," Vol. II Part 1, p.78.

$$(8)\quad D_t^2 R = D_x^2 F(x, y, z, \ldots) X^2 + D_y^2 F(x, y, z, \ldots) Y^2 \ldots$$
$$+ 2 D_x D_y F(x, y, z, \ldots) X Y + \ldots$$
$$+ D_x F(x, y, z, \ldots) D_t X + D_y F(x, y, z, \ldots) D_t Y + \ldots$$
$$\ldots\ldots\ldots\ldots$$

so that the general value of $D_t^n R$ will be composed of terms which will be the product of a positive integer, one of the partial derivatives of different order of the function $F(x, y, z, \ldots)$, and the powers of the functions $X, Y, Z, \ldots$ or their derivatives. Next, let

$$x', y', z' \ldots, t'$$

be the absolute values of the imaginary increments assigned to the variable quantities

$$x, y, z, \ldots t$$

and selected in such a way that for these absolute values, and for lesser absolute values, the functions

$$X, Y, Z, \ldots, F(x, y, z, \ldots),$$

modified by these increments, will remain continuous with respect to the arguments and the absolute values of the increments with which we deal. Finally, let

$$X', Y', Z', \ldots, R'$$

be the maxima of the absolute values of the functions

$$X, Y, Z, \ldots, R = F(x, y, z, \ldots)$$

corresponding to the absolute values

$$x', y', z', \ldots, t'$$

of the imaginary increments assigned to the variables $x, y, z, \ldots t$.

From the theorem established in a preceding section[1]

for obtaining the upper limits,

$$\mathcal{J}_0,\ \mathcal{J}_1,\ \mathcal{J}_2,\ \ldots,$$

respectively, the upper absolute values of the quantities

$$R,\quad \frac{1}{1}D_t R,\quad \frac{1}{1\cdot 2}D_t^2 R,\ldots$$

it is sufficient to calculate these quantities in the special case in which one has

$$(9)\qquad \begin{aligned} X &= ax^{-1}\,y^{-1}z^{-1}\ldots t^{-1}, \\ Y &= bx^{-1}\,y^{-1}z^{-1}\ldots t^{-1}, \\ Z &= cx^{-1}\,y^{-1}z^{-1}\ldots t^{-1}, \\ &\ldots\ldots\ldots\ldots \end{aligned}$$

and

$$(10)\qquad R = Kx^{-1}y^{-1}z^{-1}\ldots,$$

a, b, c, ... designating constant factors, and then to assign to the variables $x, y, z, \ldots, t$ and the constants $a, b, c, \ldots, K$ the values determined by this system of formulas

$$(11)\qquad \begin{matrix} x = -x', & y = -y', & z = -z', \ldots, & t = -t', \\ X = X' & Y = Y' & Z = Z' \ldots & R = R' \end{matrix}$$

related to equations (9) and (10). For the rest, to deduce the sequence (7) from the sequence (6) it will be sufficient to join with the formulas (11) the following

$$(12)\qquad |\tau - t| = \iota;$$

and, in the special case, which we have to consider, the sequence (6)

[1] "Complete Works of Cauchy", S. I. T. VI, p. 464--Extrait no.167.

will not cease to represent the expansion of

$$F(\xi, \eta, \zeta, \ldots)$$

corresponding to the values $\xi, \eta, \zeta \ldots$ which give the integration of equations (1). Finally, if the absolute value $i_g\,\tau - t$ is sufficiently small so that sequence (7) is convergent, the sequence (6) will be even more convergent. Therefore, to establish the existence of general integrals of equations (1), and furthermore, to obtain a limit within which the difference $\tau - t$ can vary without the integrals ceasing to be developable into convergent series of integer powers of this difference, it will be sufficient to integrate the system of auxiliary equations

$$(13) \qquad \begin{aligned} D_t x &= a x^{-1} y^{-1} z^{-1} \ldots t^{-1}, \\ D_t y &= b x^{-1} y^{-1} z^{-1} \ldots t^{-1}, \\ D_t z &= c x^{-1} y^{-1} z^{-1} \ldots t^{-1}, \\ & \ldots \ldots \ldots \end{aligned}$$

If the functions X, Y, Z,... did not contain the variable t, then in the values of these functions determined by the formulas (9), one would evidently have to suppress the factor t^{-1}. Then the formulas (9) would become

$$(14) \qquad \begin{aligned} X &= a x^{-1} y^{-1} z^{-1} \ldots, \\ Y &= b x^{-1} y^{-1} z^{-1} \ldots, \\ Z &= c x^{-1} y^{-1} z^{-1} \ldots, \\ & \ldots \ldots \ldots \end{aligned}$$

and the equations (13) would reduce to the following

$$
(15)\qquad
\begin{aligned}
D_t x &= a x^{-1} y^{-1} z^{-1} \ldots, \\
D_t y &= b x^{-1} y^{-1} z^{-1} \ldots, \\
D_t z &= c x^{-1} y^{-1} z^{-1} \ldots, \\
&\ldots\ldots\ldots\ldots
\end{aligned}
$$

3·22. Integration of the Auxiliary Equations.

Let us consider the system of auxiliary equations

$$
(1)\qquad
\begin{aligned}
D_t x &= a x^{-1} y^{-1} z^{-1} \ldots t^{-1}, \\
D_t y &= b x^{-1} y^{-1} z^{-1} \ldots t^{-1}, \\
D_t z &= c x^{-1} y^{-1} z^{-1} \ldots t^{-1}, \\
&\ldots\ldots\ldots\ldots
\end{aligned}
$$

in which a, b, c,... designate constant quantities. One will deduce from these

$$(2)\qquad \frac{D_t x}{a} = \frac{D_t y}{b} = \frac{D_t z}{c} = \ldots;$$

then, by integrating the formulas (2) and designating by

$$\xi, \eta, \zeta, \ldots, \tau$$

a new system of values corresponding to the variable quantities

$$x, y, z, \ldots, t,$$

one will find

$$(3)\qquad \frac{x-\xi}{a} = \frac{y-\eta}{b} = \frac{z-\zeta}{c} = \ldots .$$

Now, let us represent the value common to all the ratios which constitute the different members of the formula (3) by the letter δ, or what is the same, by the ratio δ/k, k designating a new arbitrary constant.

Then the formula

$$(4)\qquad \frac{x-\xi}{a} = \frac{y-\eta}{b} = \frac{z-\zeta}{c} = \ldots = \frac{\delta}{k}$$

becomes

$$(5)\qquad x = \xi + \frac{a}{k}\delta,\quad y = \eta + \frac{b}{k}\delta,\quad z = \zeta + \frac{c}{k}\delta, \ldots;$$

and the last equations combined with the formula (2) give

$$D_t\delta = kx^{-1}y^{-1}z^{-1}\ldots t^{-1},$$

or, what amounts to the same

(6) $$dt/t = xyz\ldots d\delta/k;$$

then, integrating the two members of the formula (6), having substituted for x, y, z,... the values given by formulas (5), one will find that

(7) $$L\left(\frac{t}{\tau}\right) = \int_0^\delta \left(\xi + \frac{a}{k}\theta\right)\left(\eta + \frac{b}{k}\theta\right)\left(\zeta + \frac{c}{k}\theta\right)\cdots\frac{d\theta}{k}.$$

Thus, the integrals of the auxiliary equations are represented by the formulas (5), the value of δ being given by the formula (7). One would be able, in these formulas, to suppose the constant k equal to unity; but, to render the application easier, which is the end we seek, it would be better not to suppose k equal 1. If the auxiliary equations reduce to the following

(8) $$\begin{cases} D_t x = ax^{-1}y^{-1}z^{-1}\ldots, \\ D_t y = bx^{-1}y^{-1}z^{-1}\ldots, \\ D_t z = cx^{-1}y^{-1}z^{-1}\ldots, \\ \ldots\ldots\ldots\ldots, \end{cases}$$

then the first member of the formula (6) reduces simply to the differential dt, and, in the place of the equation (7) one would obtain

(9) $$t - \tau = \int_0^\delta \left(\xi + \frac{a}{k}\theta\right)\left(\eta + \frac{b}{k}\theta\right)\left(\zeta + \frac{c}{k}\theta\right)\cdots\frac{d\theta}{k}.$$

3.23. Consequences of the formulas established in the preceding sections.

In the particular case in which the system of given differential equations reduces to the system of auxiliary equations and where one supposes, besides, that

$$R = F(x, y, z, \ldots) = Kx^{-1}y^{-1}z^{-1}\ldots,$$

not only do we have, by virtue of formula (4), Section 3.22

(1) $\xi = x - \frac{a}{K}\delta, \quad \eta = y - \frac{b}{K}\delta, \quad \zeta = z - \frac{c}{K}\delta, \ldots,$

the value of δ being determined by formula (7) of the same section, which can be reduced to

$$L\left(\frac{t}{\tau}\right) = \int_0^\delta \left[x + \frac{a}{K}(\theta - \delta)\right]\left[y + \frac{b}{K}(\theta - \delta)\right] \ldots \frac{\delta\theta}{K},$$

or, what amounts to the same,

(2) $$L\left(\frac{t}{\tau}\right) = \int_0^\delta \left(x - \frac{a}{K}\theta\right)\left(y - \frac{b}{K}\theta\right) \ldots \frac{d\theta}{K},$$

but also

$$F(\xi, \eta, \zeta, \ldots) = K\xi^{-1}\eta^{-1}\zeta^{-1}\ldots,$$

and, hence, using formula (1),

(3) $$F(\xi, \eta, \zeta, \ldots) = K\left(x - \frac{a}{K}\delta\right)^{-1}\left(y - \frac{b}{K}\delta\right)^{-1}\left(z - \frac{c}{K}\delta\right)^{-1}\ldots.$$

That standing, let us suppose that, in the general case where the differential equations and the function F(x, y, z,...) are of any form whatever, one constructs the sequence

(4) $$R, \quad \frac{\tau - t}{1} D_t R, \quad \frac{(\tau - t)^2}{1 \cdot 2} D_t^2 R, \ldots,$$

which, according to Taylor's formula would represent the expansion of $F(\xi, \eta, \zeta \ldots)$ arranged according to the ascending powers of $\tau - t$.

In order to obtain another sequence

(5) $\qquad \mathfrak{J}_0,\ \mathfrak{J}_1 \iota,\ \mathfrak{J}_2 \iota^2,\ \ldots,$

of which the different terms are respectively greater than the absolute values of the terms of the sequence (4), it is sufficient, by virtue of the principles established in Section 3·21 to expand according to the ascending powers of ι the value S of $F(\xi, \eta, \zeta, \ldots)$ determined by the formula (3) combined with the equation (2), after having substituted for the quantities

$$x,\ y,\ z, \ldots, t,\ \tau, \text{ and } a,\ b,\ c, \ldots, \mathfrak{K}$$

their values taken from formulas (9), (10), (11), (12) of Section 3·21. Now, if one chooses the constant k so that one has

$$(-x')(-y')(-z')\ldots(-t') = -k,$$

the formulas (9), (10), (11), (12) of Section 3·21 will give, not only

$$x = -x', \quad y = -y', \quad z = -z', \ \ldots, \ t = -t',$$

and the following,

$$\frac{\tau}{t} = 1 - \frac{\iota}{t'},$$

but also

$$\frac{a}{K} = -X', \quad \frac{b}{K} = -Y', \quad \frac{c}{K} = -Z',$$
$$\mathfrak{K} = \frac{K}{t'} R'.$$

Then to obtain the sequence (5) it will be sufficient to expand, following the ascending powers of ι, the particular value S of

$$F(\xi, \eta, \zeta, \ldots\ldots)$$

determined by the system of formulas

(6) $$t'L\left(1-\frac{L}{t'}\right)^{-1}=\int_0^{\delta}\left(1-\frac{X'}{x'}\theta\right)\left(1-\frac{Y'}{y'}\theta\right)\left(1-\frac{Z'}{z'}\theta\right)\cdots d\theta,$$

(7) $$\delta = R'\left(1-\frac{X'}{x'}\delta\right)^{-1}\left(1-\frac{Y'}{y'}\delta\right)^{-1}\left(1-\frac{Z'}{z'}\delta\right)^{-1}\cdots .$$

Besides, in order that the sequence (4) be convergent, it is sufficient that the sequence (5) be convergent. One can then establish the following proposition:

3.231. Theorem I.--Given a system of differential equations with the independent variable t and the unknowns

$$x, y, z, \ldots,$$

and also the new values of the unknowns

$$\xi, \eta, \zeta, \ldots$$

which correspond to a new value τ of the variable t, one can determine by the given differential equations the values of

$$\xi, \eta, \zeta, \ldots$$

and even of

$$F(\xi, \eta, \zeta, \ldots),$$

developable into convergent series of the ascending powers of $\tau - t$, if the value of δ, determined by equation (7), joined to formula (6), is itself developable into a convergent series arranged according to ascending powers of ι.

3.232. Corollary 1. If one lets, for the sake of brevity,

(8) $$\varepsilon = t'L\left(1-\frac{L}{t'}\right)^{-1}$$

the formula (6) becomes

(9) $$\varepsilon = \int_0^{\delta}\left(1-\frac{X'}{x'}\theta\right)\left(1-\frac{Y'}{y'}\theta\right)\left(1-\frac{Z'}{z'}\theta\right)\cdots d\theta.$$

3·233. Corollary 2. If the given differential equations do not include explicitly the variable t, then, after what has been said before (see Sec. 3·22), one will be able in the formula (2) to replace the first member, that is to say the product $L\left(\frac{t}{\tau}\right)$ by the difference $t-T$; and, in establishing besides

$$(-x')\ (-y')\ (-z')\ldots = -k,$$

one will obtain in place of the formula (6) or (9), the equation

$$l = \int_0^\delta \left(1-\frac{X'}{x'}\theta\right)\left(1-\frac{Y'}{y'}\theta\right)\left(1-\frac{Z'}{z'}\theta\right)\cdots d\theta. \tag{10}$$

3·234 Corollary 3. The value of which is given by equation (8) is developable into a convergent series by the formula

$$\varepsilon = l + \frac{1}{2}\frac{l^2}{t'} + \frac{1}{3}\frac{l^3}{t'} + \cdots, \tag{11}$$

when one has

$$l < t'. \tag{12}$$

The value of δ which is determined by equation (9), is developable by the formula of Lagrange, into a convergent series arranged according to ascending powers of ε, when one has

$$\varepsilon < \int_0^t \left(1-\frac{X'}{x'}\theta\right)\left(1-\frac{Y'}{y'}\theta\right)\left(1-\frac{Z'}{z'}\theta\right)\cdots, \tag{13}$$

t being the smallest of the ratios

$$\frac{x'}{X'},\ \frac{y'}{Y'},\ \frac{z'}{Z'},\ \ldots,$$

and it suffices evidently that the conditions (12) and (13) are fulfilled in order that the value of δ, furnished by equation (7), may be expanded into a convergent series arranged in ascending powers of l. That established, one will be able to state the following proposition:

3·235. Theorem II.--The same things being assumed as in Theorem I, (3·231), the value of

$$\xi, \eta, \zeta, \ldots,$$

and even of

$$F(\xi, \eta, \zeta \qquad),$$

will be developable into a convergent series of ascending powers of $T-t$, if the absolute value ι of $T-t$ verifies simultaneously the two conditions

$$(14) \qquad \iota < t' \qquad t' L\left(1-\frac{\iota}{t'}\right) < \int_0^{\iota}\left(1-\frac{X'}{x'}\theta\right)\left(1-\frac{Y'}{y'}\theta\right)\left(1-\frac{Z'}{z'}\theta\right)\ldots d\theta.$$

Then also, in terminating after a certain number of terms, the series which represents the expansion of $F(\xi, \eta, \zeta, \ldots)$ in ascending powers of $T-t$, one obtains a remainder whose absolute value will be less than the corresponding remainder of the series which represents the expansion of S in ascending powers of ι.

3·236. Corollary 1. If the given differential equations do not include explicitly the variable t, then, the formula (10) having been replaced by formula (13), the first of the conditions (14) will disappear and the second will be replaced by

$$(15) \qquad \iota < \int_0^{\iota}\left(1-\frac{X'}{x'}\theta\right)\left(1-\frac{Y'}{y'}\theta\right)\left(1-\frac{Z'}{z'}\theta\right)\ldots d\theta.$$

In concluding this memoir, we make one important observation. The absolute values of the different terms of the expansion of $F(\xi, \eta, \zeta, \ldots)$, arranged in ascending powers of $T-t$, will not cease to be less than the absolute values of the corresponding terms of the expansion of S, if one increases the latter. Now that is exactly what one does when one substitutes for each of the ratios

$$\frac{X'}{x'}, \frac{Y'}{y'}, \frac{Z'}{z'}, \ldots,$$

the quantity equal to the maximum among them, that is to say, the positive quantity $\frac{1}{t}$, considering that, in the expansion of S, each term will be positive and proportional to a positive power of each of these ratios. It follows from this observation that the formulas (6) and (7) would be replaced by the following

$$(16) \qquad t\,l\left(1-\frac{t}{t}\right)^{-1} = \int_0^\delta \left(1-\frac{\theta}{t}\right)^{n-1} d\theta = \frac{t}{n}\left[1-\left(1-\frac{\delta}{t}\right)^n\right],$$

$$(17) \qquad S = R'\left(1-\frac{\delta}{t}\right)^{-(n-1)},$$

from which one concludes

$$(18) \qquad S = R'\left[1-n\frac{1}{t}L\left(1-\frac{t}{t'}\right)^{-1}\right],$$

n designating the total number of the variables x, y, z,...,t.

In substituting the formula (10) for the formula (6), as one can do when the given differential equations do not include explicitly the variable t, one would obtain instead of (18),

$$(19) \qquad S = R'\left(1-n\frac{1}{t}\right)^{-\frac{n-1}{n}}.$$

Similarly, the second of the conditions (14) and the condition (15) can be, if one wishes, replaced by the following:

(20) $$t' L\left(1-\frac{\iota}{t'}\right) < \frac{\iota}{n},$$

and

(21) $$\iota < \frac{\iota}{n}$$.

The formula (21), in the case in which we suppose $n = 2$, is found in the lithographed memoir of 1835.

3·4. SIMPLIFICATIONS AND DEVELOPMENTS BY OTHER SCIENTISTS.--It is easy to see that the original form of Cauchy's demonstration based on the Calculus of Limits is much more complicated than necessary. It has been greatly simplified by Briot and Bouquet[1] in France. Following is a simplification of the proof according to the plan of Briot and Bouquet.

3·41. First, let us take a single equation

(1) $$dy/dx = f(x,y).$$

Let it be assumed that $f(x,y)$ is holoporphic (analytic) in the neighborhood of x_0 and y_0. Without any loss of generality it may be assumed also that $x_0 = y_0 = 0$. The function $f(x,y)$ will then be holomorphic with respect to x and y, when x and y are respectively within the circles C and C' drawn with the points $x = 0$ and $y = 0$ as centers, with the radii a and b, and we suppose it continuous on the circumferences themselves. Let M be the maximum absolute value of $f(x,y)$ in this region.

---

[1] Briot and Bouquet, "Journal of the École Polytechnic," Vol.XXI. (1856) pp. 133-198.

Let us suppose that equation (1) has an holomorphic solution in neighborhood of $x=0$, which vanishes when $x=0$, namely

$$y=f(x)=f(x_0)+\frac{1}{1}f'(x_0)x+\frac{1}{1\cdot2}f''(x_0)x^2+\dots$$

which may be written in the form

$$y=\left(\frac{dy}{dx}\right)_0 x+\frac{1}{1\cdot2}\left(\frac{d^2y}{dx^2}\right)_0 x^2+\dots\,.$$

We can obtain by means of equation (1) itself, the values of the successive derivatives dy/dx, d^2y/dx^2, d^3y/dx^3,...for $x=0$. It is sufficient to differentiate equation (1), first one time, to have d^2y/dx^2, and to substitute, in the second term, $x=0$, $y=0$; differentiating again, we shall have d^3y/dx^3, and so on. It should be noted that if $f(x,y)$ and all of its partial derivatives, at $x=y=0$, are positive then all the derivatives of y with respect to x are positive also, their values being derived by a process of addition and multiplication (not subtraction) of positive quantities. We shall then have

$$(2)\qquad y=\left(\frac{dy}{dx}\right)_0 x+\frac{1}{1\cdot2}\left(\frac{d^2y}{dx^2}\right)_0 x^2+\dots=a_1x+a_2x^2+\dots\,.$$

It is obvious, from the manner in which the coefficients of the series in equation (2) were determined, that if the series is convergent it is the unique solution of equation (1). Therefore, the essential point in the proof consists in demonstrating that the series thus obtained converges if x has a sufficiently small value. This point established, it is clear that the function y so determined satisfies the differential equation, since the functions of x

$$dy/dx \text{ and } f(x,y)$$

have, from the way in which y has been obtained, the same value for $x=0$, and so do their corresponding derivatives of every order; they are equal then; that is, equation (1) is established.

By comparison with another series we can prove the convergence of series (2), and the idea of such a comparison forms what is really interesting and fruitful in what Cauchy called the "Calculus of Limits."

Now, let us take the function

$$F(x,y)\ \frac{M}{(1-x/a)(1-y/b)}$$

holomorphic within the same circles C and C' and whose partial derivatives,[1] all positive for $x=0$ and $y=0$, are such that

$$\text{(a)}\qquad \left|\frac{\partial^{n+p}f}{\partial x^n \partial y^p}\right|_{\substack{x=0\\y=0}} \overline{\overline{<}} \left(\frac{\partial^{n+p}F}{\partial x^n \partial y^p}\right)_{\substack{x=0\\y=0}} \quad \begin{matrix}(n=1,2,3,\ldots)\\(p=1,2,3,\ldots)\end{matrix}.$$

Let us next consider the auxiliary differential equation

$$dY/dx = F(x,Y),$$

and suppose, as we shall prove later, that there exists an integral Y of this equation, holomorphic in the neighborhood of $x=0$ and vanishing for $x=0$. We shall then have

$$\text{(3)}\qquad Y = \left(\frac{dY}{dx}\right)_0 x + \frac{1}{1\cdot 2}\left(\frac{d^2Y}{dx^2}\right)_0 x^2 + \cdots = A_1 x + A_2 x^2 + \cdots \quad .$$

The coefficients of the powers of x in this series are, for reasons mentioned above, positive, and from the inequalities (a) we shall obviously have

$$|a_m| < A_m .$$

The series (2) will then certainly be convergent in the regions in which the series (3) is convergent. Now, it is easy to show the

[1]Picard, "Traite D' Analyse," Vol. II, p. 259.

existence of the function Y. Let us write the equation

$$\frac{dY}{dx} = \frac{M}{(1-\frac{x}{a})(1-\frac{Y}{b})}$$

in the form

$$\left(1-\frac{Y}{b}\right)\frac{dY}{dx} = \frac{M}{1-\frac{x}{a}} \, .$$

If the function Y exists, the two members are respectively the derivatives of

$$Y - Y^2/2b \quad \text{and} \quad - Ma \log(1-x/a)$$

We will take the value of the logarithm which vanishes for $x = 0$,

$$\log(1-x/a) = -x/a - x^2/2a^2 - x^3/3a^3 \quad \ldots.$$

Since Y vanishes for $x = 0$, we shall then have

$$Y - \frac{Y}{2b} = - Ma \log\left(1-\frac{x}{a}\right);$$

and, consequently,

$$Y = b - b\sqrt{1 + \frac{2Ma}{b}\log\left(1-\frac{x}{a}\right)}$$

taking for the radical the value $+1$ for $x = 0$.

The function Y, thus determined, satisfies the equation

$$dY/dx = F(x,Y).$$

It vanishes for $x = 0$, and it is holomorphic within a circle having as center the origin and a radius r which causes the quantity under the radical to vanish; that is, r is determined by the equation

$$1 + \frac{2M}{1}\frac{a}{b}\log\left(1-\frac{r}{a}\right) = 0$$

which gives

$$r = a\left(1 - e^{-\frac{b}{2Ma}}\right).$$

We are then certain that the series (3) converges within the circle of radius r; it is likewise true then that series (2) converges in the same region, and consequently, we may state that equation (1) has an holomorphic integral in the circle of radius r having the origin as center, that it vanishes for $x=0$, and that it is unique.

It may be noticed that within the circle of radius r we have

$$|Y| < b;$$

and consequently

$$|y| < b$$

within the same circle.

3·42. Application to n equations. The preceding analysis covers without modification the case of n equations

$$dy_1/dx = f_1(x, y_1, y_2, \ldots, y_n),$$
$$dy_2/dx = f_2(x, y_1, y_2, \ldots, y_n),$$
$$\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot,$$
$$dy_n/dx = f_n(x, y_1, y_2, \ldots, y_n).$$

We suppose that the functions $f_1, f_2, \ldots, f_n$ are holomorphic with respect to x and y in the circles of radius a and b, respectively, having the origin as center in the planes of x and y. Let M denote the maximum absolute value of the f's in the designated regions and let us compare this system with the following

$$dY_1/dx = dY_2/dx = \ldots = dY_n/dx = F(x, Y_1, Y_2, \ldots, Y_n),$$

taking

$$\text{(a)} \quad F(x, Y_1, Y_2, \ldots, Y_n) = \frac{M}{(1-x/a)(1-Y_1/b)\ldots(1-Y_n/b)}.$$

Since the Y's must vanish for $x=0$ and the second member of the equation (a) is symmetrical with respect to the Y's, the Y's are identical and we only have to consider the single equation

$$dY/dx = \frac{M}{(1-x/a)(1-Y/b)^n} .$$

The radius r of the circle in which the series are known to be convergent, will, in this case, be

$$r = a\left(1 - e^{-\frac{b}{(n+1)Ma}}\right).$$

3·43. Cauchy's second proof has also been modified in Germany by Weierstrass (1815-1897)[1].

Following is the essence of Weierstrass' exposition.

Suppose that the differential equation $F(y',y,x) = 0$ (where y' stands for dy/dx) is put in the form

(1) $$y' = f(x,y)$$

which is always possible. The proof is limited to the case where $f(x,y)$ is a function which can be represented by a power series

$$a_{oo} + a_{1o}x + a_{o1}y + a_{2o}x^2 + a_{11}xy + a_{o2}y^2 + \cdots$$

in which the a's are all known, since $f(x,y)$ is known, and which converges for $|x| \leqq r$, $|y| \leqq t$, say. Without any loss of generality $y = 0$ when $x = 0$ is taken for the initial condition.

It is to be proved (a) that there is <u>one</u> and <u>only one</u> series

(2) $$y = b_1x + b_2x^2 + b_3x^3 + \cdots$$

which identically satisfies

(3) $$y' = a_{oo} + a_{1o}x + a_{o1}y + a_{11}xy + a_{2o}x^2 + a_{o2}y^2 + \cdots + a_{mn}x^my^n + \cdots,$$

and (b) that within certain limits for x this series is convergent.

---

[1] Weierstrass, "Mathematical Werke," Vol. I, pp.67-75, (1842).

On transforming the series in (3) which has been supposed convergent for $|x| \leq r, |y| \leq t$, by putting $x = rx_1$, $y = ty_1$, equation (3) takes the form

$$y' = f(rx_1, ty_1) = a'_{oo} + a'_{1o}x_1 + a'_{o1}y_1 + a'_{2o}x_1^2 + a'_{11}x_1y_1 + a'_{o2}y_1^2 + \dots$$

The second member of this equation is a convergent series, and converges when $x_1 = y_1 = 1$; and, therefore, $a'_{oo} + a'_{1o} + a'_{o1} + \dots$ converges. This shows that the absolute value of each a' is not larger than a certain finite quantity, say A. The substitution just made for x and y does not make any essential change in the problem, and hence, it might have been assumed at first that the a's of (3) were each not greater than A. In what follows, therefore, the a's are regarded as not greater than A.

If (2) satisfies (3), the value of y and y' derived from (2), when substituted in (3), must make the latter an identity; and, therefore,

$$b_1 + 2b_2x^1 + 3b_3x^2 + 4b_4x^3 + \dots$$
$$= a_{oo} + a_{1o}x + a_{o1}(b_1x + b_2x^2 + \dots) + a_{2o}x^2 + a_{11}x(b_1x + b_2x^2 + b_3x^3 + \dots)$$
$$+ a_{o2}(b_1x + b_2x^2 + b_3x^3 \quad \dots)$$

is an identical equation. Hence,

$$b = a_o; \; 2b_2 = a_{1o} + a_{o1}b_1,$$

that is

$$b_2 = \tfrac{1}{2}(a_{10} + a_{01}a_{00}); \; b_3 = a_{20} + \frac{a_{01}}{2}(a_{10} + a_{01}a_{00}) + a_{11}a_{00};$$

and likewise for $b_4, b_5, \dots$. It is evident that all the b's can be determined as rational integral functions of the a's; and also that all the numerical coefficients in the expressions for the b's are positive; and, consequently, the b's will not be diminished if each of the a's is replaced by A.

From the method of derivation, it is evident that (2), with the b's determined as above, identically satisfies (3) and that it is unique. It must now be determined whether this series is convergent.

On replacing each of the a's in (3) by A, a quantity not less than any of the a's, there results

(4) $y' = A(1 + x + y + x^2 + xy + y^2 + x^3 + x^2 y + \dots)$.

The integral of this equation is found by replacing each of the a's that occur in the expressions for the b's of (2) by A. None of these latter coefficients is diminished by changing each of the a's to A, as pointed out above, consequently, if the integral of (4) is convergent, the integral of (3) is also.

Now, let us solve (4) directly. On factoring the second member, the equation becomes

$$y' = A(1 + x + x^2 + \dots)(1 + y + y^2 + y^3 + \dots) = A\frac{1}{1-x}\cdot\frac{1}{1-y}.$$

Therefore,

$$(1-y)dy = A\frac{dx}{1-x},$$

and upon integrating we get $y - \frac{1}{2}y^2 = -A\log(1-x) + c$.

We take for the logarithm that branch which vanishes for $x = 0$, and from the initial condition that $y = 0$ when $x = 0$ we see that $c = 0$.

Consequently, $y = 1 - [2A\log(1-x) + 1]^{\frac{1}{2}}$.

Since $y = 0$ when $x = 0$ we must take for

$$[2A\log(1-x) + 1]^{\frac{1}{2}}$$

that branch which, for $x = 0$, has the value $+1$.

Then

(5) $y = 1 - [1 + 2A\log(1-x)]^{\frac{1}{2}} = 1 - [1 - 2A(x + \frac{x^2}{2} + \frac{x^3}{3} + \dots)]^{\frac{1}{2}}$.

The series

$$x + \frac{x^2}{2} + \frac{x^3}{3} + \cdots$$

converges when the absolute value of x is less than 1; consequently the value of y in (5) is finite, and, therefore, the value of y in (2) is finite for x within certain limits. The proof for n equations can be carried out in much the same way.

Note: In Sec. 5·3 are listed the names and works of other mathematicians who have given good treatments of the Calculus of Limits or some of its phases.

CHAPTER IV

"THE METHOD OF SUCCESSIVE APPROXIMATIONS"

4·1. DEVELOPED BY PICARD.--The method of successive approximations was probably known to Cauchy but appears to have been first published by Liouville, "Journal de Mathematics" (1) 2 (1837), p. 19; (1) 3 (1838), p. 565, who applied it to the case of the homogeneous linear equation of the second order. Extensions to the linear equation of order n are given by J. Caque, "Journal of Mathematics" (2) 9 (1864), p. 185; L. Fuchs, "Annali di Math." (2) 4 (1870), p. 36 (Ges. Werke, I. p.295); and G. Peano, "Math. Ann." 32 (1888), p. 450.[1]

In its most general form, however, it has been developed by Charles Emile Picard (1856-   ) of France.[1]

Picard was born in Paris and was educated at The École Normale Superieure where he was inspired by J.G. Darboux. In 1877 he was lecturer on mathematics in Paris. From 1879 to 1881 he was lecturer on the same subject at Toulouse. He married a daughter of Hermite in 1881, and that same year returned to Paris and became Professor of Mathematics at The École Normale and the Sorborne.

His most famous work is his "Traite d' Analyse" (3 Vols., 1891-96; 2nd Ed. 1901-8) which is still a standard textbook. His other works include "Theorie des Fonctions Algebraques de deux Variables Independantes" (1897-1906), with G. Simart; and "Sur le Developpement de l' Analyses et Ses Rapport Avec Diverse Sciences" (1905), lectures delivered in America.

---

[1]Ince, "Ordinary Differential Equations" (1927) p. 63.

Picard's demonstration of the existence theorem is remarkable for its simplicity and brevity, and requires no auxiliary propositions. It is so simple and complete that no other mathematician has been able to improve it. Picard's proof was first given to the public in 1890, while he was lecturing in Paris, in his memoir published in "Journal de Mathematique." In the following year it was published in the "Bulletin de la Societe Mathematique de France" for March, and was reproduced, on account of its striking character, in the "Nouvelles Annales des Mathematiques" for May. A translation of Picard's proof by T. S. Fiske was published in "Bulletin of the New York Mathematical Society" Vol. I, pp. 12-16 (1891-92). The most complete form of Picard's proof may be found in his "Traite d' Analyse." The translation by Fiske was made during the early stages of Picard's work on this theorem and, therefore, it is not the most general form of the proof. For that reason we are presenting in the following section a translation of the demonstration as it is stated in "Traite d' Analyse" (2nd Ed.) Vol. II pp.340-344.

4·2. A TRANSLATION OF PICARD'S DEMONSTRATION.---Consider the system of n equations of the first order,

$$
\begin{aligned}
\frac{du}{dx} &= f_1(x, u, v, \cdots, w), \\
\frac{dv}{dx} &= f_2(x, u, v, \cdots, w), \\
&\cdots\cdots\cdots \\
\frac{dw}{dx} &= f_n(x, u, v, \cdots, w)
\end{aligned}
$$

in which the functions $(f_1, f_2, f_3, \ldots, f_n)$ are continuous real functions of the real quantities x, u, v,..., w in the neighborhood of x_0, $u_0, \ldots, w_0$, and have determinate values as long as x, u, v,...,w remain within the respective intervals

$$(x_0 - a,\ x_0 + a),$$
$$(u_0 - b,\ u_0 + b),$$
$$(v_0 - b,\ v_0 + b),$$
$$\cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot$$
$$(w_0 - b,\ w_0 + b),$$

a and b denoting two positive magnitudes.

Suppose that n positive quantities

$$A,\ B, \ldots,\ L$$

can be determined in such a manner that

$$\left| f(x,\ u',\ v', \ldots, w') - f(x,\ v,\ u, \ldots, w) \right| < A\left|u' - u\right| + B\left|v' - v\right| + \cdots + L\left|w' - w\right|,$$

in which $|\alpha|$ denotes as usual the absolute value of α and x, v, u,...,w are contained in the indicated intervals. This will evidently be the case when the functions have finite partial derivatives with respect to u, v,...,w.

Starting with these very general hypotheses we will demonstrate that <u>there exist functions u, v,...,w of x, continuous in the neighborhood of x_0, satisfying the given differential equations, and reducing, for x x_0, respectively to u_0, v_0,...,w_0.</u>

We proceed by successive approximations. Taking first the system

$$\frac{du_1}{dx} = f_1(x, u_0, v_0, \ldots, w_0),$$
$$\frac{dv_1}{dx} = f_2(x, u_0, v_0, \ldots, w_0),$$
$$\cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot$$
$$\frac{dw_1}{dx} = f_n(x, u_0, v_0, \ldots, w_0),$$

we obtain by quadratures the functions

$$u_1, v_1, \dots, w_1,$$

determining them in such a manner that they take for x_0 the values

$$u_0, v_0, \dots, w_0.$$

Forming then the equations

$$\frac{du_2}{dx} = f_1(x, u_1, v_1, \dots, w_1),$$
$$\frac{dv_2}{dx} = f_2(x, u_1, v_1, \dots, w_1),$$
$$\dots\dots\dots$$
$$\frac{dw_2}{dx} = f_n(x, u_1, v_1, \dots, w_1),$$

we determine $u_2, v_2, \dots, w_2$, by the conditions that they take for x_0 the values $u_0, v_0, \dots, w_0$, respectively. We continue this process indefinitely, the functions $u_m, v_m, \dots, w_m$ being connected with the preceding $u_{m-1}, v_{m-1}, \dots, w_{m-1}$ by the relations

$$\frac{du_m}{dx} = f_1(x, u_{m-1}, v_{m-1}, \dots, w_{m-1})$$
$$\dots\dots\dots$$
$$\frac{dw_m}{dx} = f_n(x, u_{m-1}, v_{m-1}, \dots, w_{m-1})$$

and, for $x = x_0$, satisfying the equations

$$u_m = u_0, \quad v_m = v_0, \dots, w_m = w_0.$$

We will now prove that when m increases indefinitely, $u_m, v_m, \dots, w_m$ tend toward limits which represent the integrals sought provided x remains sufficiently near x_0.

Let M be the maximum absolute value of the functions f_i $(i=1,2,3, \dots, n)$ when the variables upon which they depend remain between the indicated limits. Denote by r a quantity at most equal to a. If now x remains in the interval

$$(x_0 - r), \quad (x_0 + r)$$

we have

$$|u_1 - u_0| \leqq M|x - x_0|, \ldots, |w_1 - w_0| \leqq M|x - x_0|.$$

Hence, provided $M \cdot r \leq b$, the quantities u_1, $v_1, \ldots,$ w_1 remain within the desired limits, and it is evident that the same is true of all the other sets of values of u_m, $v_m, \ldots,$ w_m. Denoting by d a quantity at most equal to r, suppose that x remains in the interval

$$(x_0 - d,\ x_0 + d),$$

and write

$$u_m - u_{m-1} = U_m, \ldots, w_m - w_{m-1} = W_m \,.$$

We have, placing $m = 2, 3, \ldots,$ that

$$\frac{dU_m}{dx} = f_1(x, u_{m-1}, v_{m-1}, \ldots, w_{m-1}) - f_1(x, u_{m-2}, \ldots, w_{m-2}),$$

$$\cdots\cdots\cdots\cdots\cdots$$

$$\frac{dW_m}{dx} = f_n(x, u_{m-1}, \ldots, w_{m-1}) - f_n(x, u_{m-2}, \ldots, w_{m-2}).$$

Since

$$|U_1| \leqq M|x - x_0|, \ \ldots, \ |W_1| \leqq |x - x_0| M,$$

the preceding equations show, by means of the Lipschitz condition, that

$$\frac{d|U_m|}{dx} \leqq A|U_{m-1}| + B|V_{m-1}| + \cdots + L|W_{m-1}|,$$

$$\frac{d|V_m|}{dx} \leqq A|U_{m-1}| + B|V_{m-1}| + \cdots + L|W_{m-1}|,$$

$$\cdots\cdots\cdots\cdots\cdots$$

$$\frac{d|W_m|}{dx} \leqq A|U_{m-1}| + B|V_{m-1}| + \cdots + L|W_{m-1}|.$$

In particular, when m 2 we get, by using

$$|U_1| \leqq M|x-x_0|, \; \ldots, \; |W_1| \leqq |x-x_0|,$$

the following:

$$\left|\frac{dU_2}{dx}\right| \leqq M(A+B+\cdots+L)|x-x_0|,$$

. . . .

$$\left|\frac{dW_2}{dx}\right| \leqq M(A+B+\cdots+L)|x-x_0|,$$

By integrating we get

$$|U_2| \leqq \left|\int_{x_0}^{x} |x-x_0| M(A+B+\cdots+L)\,dx\right|,$$

. . . .

$$|W_2| \leqq \left|\int_{x_0}^{x} |x-x_0| M(A+B+\cdots+L)\,dx\right|,$$

or

$$|U_2| \leqq M(A+B+\cdots+L)\frac{|x-x_0|^2}{1\cdot 2},$$

and the same limit holds for

$$|V_2|, \ldots, |W_2|.$$

And in like manner we get, for m = 3,

$$|U_3| \leqq \left|\int_{x_0}^{x} M(A+B+\cdots+L)^2\frac{|x-x_0|^2}{1\cdot 2}\,dx\right| = M(A+B+\cdots+L)^2\frac{|x-x_0|^3}{1\cdot 2\cdot 3}.$$

And the same limit holds for

$$|V_3|, \ldots, |W_3|.$$

Continuing step by step, it may be shown that

$$|U_n|, \ldots, |W_n| \leqq M(A+B+\cdots+L)^{n-1}\frac{|x-x_0|^n}{1\cdot 2\cdots n}.$$

Therefore, each of the series

$$u = u_0 + (u_1 - u_0) + \dots + (u_n - u_{n-1}) + \dots$$

$$v = v_0 + (v_1 - v_0) + \dots + (v_n \quad v_{n-1}) + \dots$$

$$\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot$$

$$w = w_0 + (w_1 - w_0) + \dots + (w_n \quad w_{n-1}) + \dots$$

is uniformly convergent in the interval

$$(x_0 - d, \quad x_0 + d),$$

d being the smaller of the two quantities a, b/M.

Moreover, we have

$$u_n = \int_{x_0}^{x} f_1(x, u_{n-1}, v_{n-1}, \dots, w_{n-1})\, dx + u_0,$$

and, since u_n, $v_n, \dots$, w_n uniformly approach the limits u, $v, \dots$, w, we have in the limit

$$u = \int_{x_0}^{x} f(x, u, v, \dots, w)\, dx + u_0;$$

hence

$$\frac{du}{dx} = f_1(x, u, v, \dots, w).$$

Similar results held for the other functions.

The functions u, $v, \dots$, w, are therefore the integrals sought. They are defined in the interval $(x_0 - d, x_0 + d)$.

Note: No important contributions to the method of successive approximations have been made since it was developed by Picard. Some observations and references on the variation and necessity of the conditions involved will be found in the first part of Chapter V. Ince, "Ordinary Differential Equations" (1926) pp. 63-66 and Goursat-Hedrick-Dunkel, "Mathematical Analysis," pp. 61-68, are among the modern writers who have given good treatments of this method.

CHAPTER V

GENERAL OBSERVATIONS

5·1. OBSERVATIONS ON THE METHOD OF SUCCESSIVE APPROXIMATIONS.-- Ernst Lindelof[1] has demonstrated that, in a great many cases, there exists an interval more extended than $(x_0 - a,\ x_0 + a)$ in which the integrals are continuous. If the functions $(f_1, f_2, f_3, \ldots, f_n)$ are continuous for all the values of x in the interval $(x_0 - a,\ x_0 + a)$ and for all values of $u, v, \ldots, w$, then it is unnecessary to make the requirement that $M \cdot a < b$ in the proof of Picard. In order to prove the convergence of the series expressing the values of the integrals $(u, v, \ldots, w)$, it is sufficient that there exists positive quantities $(A, B, \ldots, L)$ such that the Lipschitz condition remains true for all values $u, v, \ldots w$ when x remains in the interval $(x_0 - a,\ x_0 + a)$. [2]According to the law of the mean, these conditions are satisfied if the functions $(f_1, f_2, \ldots, f_n)$ have partial derivatives with respect to the variables $u, v, \ldots, w$ which remain finite for all values of $u, v, \ldots, w$ when x remains in the interval $(x_0 - a,\ x_0 + a)$.

Ince[3] has also made some interesting observations on the method of successive approximations of which a brief summary is here given. He notes that the continuity of the functions $(f_1, f_2, \ldots, f_n)$ is not necessary for the existence of continuous solutions. They may admit

---

[1]E. Lindelof, "Journal de Mathematiques," (4) 10 (1894), p.117; "C. R. Acad. Sc. Paris," 118 (1894) p. 454.

[2]Goursat-Hendrick-Dunkel, "Mathematical Analysis" Vol.II, Part II, p. 65.

[3]Ince, "Ordinary Differential Equations" (1927) pp. 66-68.

of a limited number of finite discontinuities, consisting of discrete points or lines parallel to the axes of the dependent variables. Any other lines of discontinuity imply a violation of the Lipschitz condition. Mie, "Mathematical Annals" 43 (1893) p. 553, has shown that solutions exist whenever the functions are continuous in the dependent variables and discontinuous but integrable (in Riemann's sense) with respect to the independent variable.[1]

If the functions (f_1, f_2,..., f_n) are continuous but do not satisfy the Lipschitz condition, I. Bendixson[2] has demonstrated that if the approximation series of Picard converge, the limits of their convergence will be solutions of the given equations.

The Lipschitz condition, or a condition of similar character, is necessary to insure the uniqueness of the solutions. Peano[3] and Osgood[4] have proved that, if $f(x,y)$ be continuous in the neighborhood of (x_o, y_o), and the Lipschitz condition, or a similar one, is not imposed, there exists in general a one-fold infinity of solutions satisfying the initial conditions. However, the uniqueness of the solution is not destroyed if the Lipschitz condition is replaced by

---

[1]The differential equations are transformed into integral equations. See Bocher, "Intro. to the Theory of Integral Equations;" Whittaker and Watson, "Modern Analysis," Chap. XI.

[2]Bendixson, "Ofversigt Vetensk. Akad. forhandl." (Stockholm) 54 (1897), p. 617.

[3]Peano, "Math. Ann." 37 (1890) p. 182; "Atti. Accad. Torino," 26 (1890-91), p. 677.

[4]Osgood, "Monatsh. Math. Phys." 9 (1898), p. 331.

one or another of the less restrictive conditions

$$|f(x,Y)-f(x,y)| < K_1|Y-y| \log \frac{1}{|Y-y|},$$
$$|f(x,Y)-f(x,y)| < K_2|Y-y| \log \frac{1}{|Y-y|} \log \log \frac{1}{|Y-y|},$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$

in which K_1, K_2,..., K are constants. The value of the constants occuring in the Lipschitz condition determines the utility of the approximation series in determining the value of the integrals. Practical methods of approximate computation based upon the method of successive approximations have been devised by Severini, "Rend. 1st. Lombard." (2) 31 (1898), pp. 657, 950; Cotton, "C. R. Acad. Sc. Paris," 140 (1905), p. 494; 141 (1905), p. 177; 146 (1908), pp. 274, 510; "Math. Ann.," 31 (1908), p. 107.

5·2. OBSERVATIONS ON THE METHOD OF DIFFERENCE EQUATIONS.--Among the French mathematicians the most noted treatment of the method of difference equations was given by Louis Philippe Gilbert, "Cours de mecanique Analytique," (1877).

V. Volterra[1] has given a new demonstration of this method, which enlarges a little the condition of Lipschitz. The method of V. Volterra has been interpreted graphically by G. Picciati[2].

Painleve[3] has proved that the interval of convergence in the Cauchy-Lipschitz method is greater than the interval $|x-x_0| < l$ in which l is the minimum of a and b/M. He found that l can be replaced

---

[1]Volterra, "Giorn. mat." (1) 19 (1881), p. 333.
[2]Picciati, "Il Politecnico" 41 (1893), p. 493, 537.
[3]Painleve, "Bull. Soc. math. France" 27 (1899), p. 149.

by the number λ equal the minimum of the two quantities

(1) a and $1/k_1+\dots+k_n \cdot \log\left[1 + b(k_1+\dots+k_n)/M_o\right]$,

where M_o designates the maximum of the absolute values of the n functions

$$f_1(x, y_1, \dots, y_n), \dots, f_n(x, y_1, \dots, y_n)$$

for $|x - x_o| \leqq a$.

In a great number of cases $\lambda > l$. If, for instance, the functions f_i $(i = 1, 2, \dots, n)$ are continuous and their derivatives are also continuous functions for $|x - x_o| < a$, whatever the values of $y_1, \dots, y_n$ may be, and if their derivatives $\frac{\partial f_i}{\partial y_j}$ remain absolutely less than a fixed quantity A, the second of the two quantities (1) will become infinite with b, and the solution corresponding to the initial conditions is continuous and unique in the whole interval $|x - x_o| < a$.[1]

Theoretically, the method of difference equations is superior to the method of successive approximations because it not only gives the interval in which the integral certainly exists, but also leads to a solution which converges uniformly throughout any greater interval $(x_o, x_o + k)$ in which the solution, defined by the assigned initial conditions, is continuous. The proof of this point may be found in Ince, "Ordinary Differential Equations" (1927) pp. 80, 81, or in Goursat-Hedrick-Dunkel, "Mathematical Analysis" Vol. II, Part II, (1917) pp. 73, 74, in which works are found two of the best modern treatments of the method of difference equations.

---

[1] Bendixson, "Ibid" 54(1897), p. 617.

This method also applies to complex variables. The investigations of E. Picard and of Painlevé have shown that it leads to developments of the integrals in convergent series in the whole region of their existence if the right-hand sides of the given equations remain analytic in this region.[1]

5·3. OBSERVATIONS ON THE METHOD CALLED "THE CALCULUS OF LIMITS."-- Some authors use the terms "holomorphic" and "analytic" interchangeably and others make a distinction. In this thesis they are used interchangeably and according to Cauchy's definition. That is, $f(z,w)$ is an analytic function of z and w in a domain D if (1) $f(z,w)$ is a continuous function of z and w in D; and (2) if $\frac{\partial f}{\partial z}$, $\frac{\partial f}{\partial w}$ both have a finite existence at every point of D.[2] The investigations of Goursat[3] have shown that, when the functions are assumed to be analytic, the method of successive approximations can be applied to the complex domain with merely verbal alterations. The development in power series of the integral is identical with that furnished by the calculus of limits, but the limit obtained for the radius of convergence is greater.[4] As stated above, (5·2), the method of difference equations can also be applied to the complex domain. However, the method of limits is the one that is perhaps most appropriate in this case.[5] The condition of

---

[1]Goursat-Hedrick-Dunkel, "Mathematical Analysis" (1917) Vol. II, Part II, p. 74; Painlevé, "C. R. Acad. Sc. Paris," 128 (1899), p. 1505; Picard, "Traite d' Analyse," 2, p. 1363; "Ann. Ec. Norm." (3), 21 (1904), p. 56.

[2]E. Picard, "Traite d' Analyse," 2, Chap. IX.

[3]Goursat-Hedrick-Dunkel, "Ibid" pp. 66, 67.

[4]Painlevé, "Bull. Soc. Math. France," 27 (1899) p. 152.

[5]Ince, "Ibid" p. 281.

analyticity when the variables are complex, replaces the condition that, when the variables are real, f is a continuous function and satisfies the Lipschitz condition. The fact that, when f(z,w) is analytic, $\frac{\partial f}{\partial \omega}$ is bounded takes the place of the Lipschitz condition in the proof of the existence of a solution.[1]

The fundamental idea of The Calculus of Limits consists in the use of dominant functions.[2] Since every analytic function has an infinite number of dominant functions, we see that the method can be varied in a great many ways. The simplicity of the demonstrations depends largely on the choice of the dominant functions. Since the work of Cauchy, his proofs have been perfected and extended to more general cases by Ch. A. A. Briot and J. C. Bouquet (3·41), K. Weierstrass (3·42), Koenigsberger,[3] Ch. Meray,[4] Riquier,[5] Madame Kovalevsky,[6] Jordan[7] and others. No improvement or change has been made in the fundamental principle of the proof. Even today the same method is constantly used to treat analogous questions relative to partial differential equations with various initial conditions.[8] Among the modern works on

---

[1]Ince, "Ibid" p. 281.

[2]For a definition of dominant functions see Goursat-Hedrick, "Ibid" Vol. I, p. 386.

[3]The work of Weierstrass (3·42) was simplified by Koenigsberger, "J. fur Math.," 104 (1889), p. 174; "lehrbuch," p. 25.

[4]Meray, "Lecons nouvelles sur l'analyse infinitesimale" 1, Paris, 1894.

[5]Riquier, "Sur les systemes d'equations aux derivees partielles".

[6]Kovalevsky, "J. de Crelle" Vol. LXXX.

[7]Jordan, "Cours d'Analyse de l'Ecole Polytechnique" Vol. I-III, (1882-87).

the subject, two of the best treatments of The Calculus of Limits that we can find are Ince, "Ordinary Differential Equations," (1927) pp. 281-284; and Goursat-Kedrick-Dunkel, "Mathematical Analysis," Vol. II, Part II, pp. 45-61.

5·4 APPLICATION OF THE EXISTENCE THEOREM TO AN EQUATION NOT OF THE FIRST DEGREE.[1]--Consider the differential equation of the form

$$F(x,y, dy/dx) = 0$$

in which F is a polynomial in dy/dx, and is single-valued in x and y. Let (x_0,y_0) be any initial pair of values of (x,y). Then, if the equation

$$F(x,y,p) = 0$$

has a non-repeated root $p = p_0$ when $x = x_0$, $y = y_0$, it will have one and only one root

$$p = f(x,y),$$

which reduces to p_0 when $x = x_0$, $y = y_0$, and f(x,y) will be single-valued in the neighborhood of (x_0,y_0).

Now, if f(x,y) is continuous and satisfies a Lipschitz condition throughout a rectangle surrounding the point (x_0,y_0), the equation

$$dy/dx = f(x,y)$$

will possess a unique solution, continuous for values of x sufficiently near to x_0, and satisfying the assigned initial conditions. This solution clearly satisfies the original equation for the same range of values of x, and thus in this case the problem presents no new features.

---

[1]Ince, "Ordinary Differential Equations" (1927) pp. 82,83.

On the other hand, when the given equation

$$F(x,y,p) = 0$$

has a multiple root $p = p_0$ for $x = x_0$, $y = y_0$, then p is a non-uniform function of (x,y) in any domain including the point (x_0,y_0) and, therefore, the existence theorem is not applicable.

5.5. DIFFERENTIAL EQUATIONS OF ORDER HIGHER THAN THE FIRST.--A single differential equation of order n, with one dependent variable, is reducible to a system of n equations of first order. Likewise, a system of n equations of any order in n dependent variables may be replaced by a system of equations of the first order by letting each of the derivatives of the dependent variables except the highest ordered, in the case of each variable, be a new variable.[1] Therefore, the existence theorem is applicable to ordinary differential equations of any order.

---

[1] Cohen, "Differential Equations" (1906) pp. 160, 161.

THE END

BIBLIOGRAPHY

Bendixson, "Ofversigt Vetensk. Akad. forhandl," 54 (1897).

Bocher, "Intro to the Theory of Integral Equations."

Briot and Bouquet, "Journal of the Ecole Polytechnic," Vol. XXI (1856).

Caque, "Journal of Mathematics," (2) 9 (1864).

Cauchy, "Collected Works," Series I, Vols. 4-7, 10; "Comptes Rendus Academie de Science Paris," 9-11, 14, 15, 23 (1839-46); "Sur l'integration des Equations Differentielles," (1835); "Exercises d'Analyse," (1840).

Cohen, "Differential Equations," (1906).

Cotton, "C. R. Acad. Sc. Paris," 140 (1905); 141 (1905); 146 (1908); "Math. Ann.," 31 (1908).

Fiske, "Bulletin of the New York Mathematical Society," Vol.I (1891-92).

Fuchs, "Annali di Math." (2) 4 (1870).

Gilbert, "Cours de mecanique Analytique," (1877).

Goursat-Hedrick-Dunkel, "Mathematical Analysis," Vol. II, Parts I-II.

Ince, "Ordinary Differential Equations" (1927).

Jordan, "Cours D'Analyse de l'Ecole Polytechnique" Vol. I-III (1882-7).

Koenigsberger, "J. fur Math." 104 (1889); "lehrbuch."

Lindelof, "Journal de Mathematiques," (4) 10 (1894); "C. R. Acad. Sc. Paris," 118 (1894).

Liouville, "Journal de Mathematiques" (1) 2 (1838); (1) 3 (1838).

Lipschitz, "Bulletin des Sciences Mathematique et Astronomiques," 10 (1876).

Madame Kovalevsky, "Journal de Crelle," Vol. LXXX.

Meray, "Lecons Nouvelles sur l'Analyse Infinitesimale."

Mie, "Mathematical Annals," 43 (1893).

Moigno, "Lecons de Calculus," Vol. II.

Osgood, "Monatsh. Math. Phys.," 9 (1898).

Painleve, "Bull. Soc. Math. France," 27 (1899); "Lecons sur la theorie Analytique des equations differentielles," (1897); "C. R. Acad. Sc. Paris," 128 (1899).

Peano, "Math. Ann." 32 (1888).

Picard, "Traite d' Analyse," Vols. I-II; "Theorie des Fonctions Algebrique de deux Variables Independentes," with G. Smart; "Sur le Developpement de l'Analysis et Ses Rapport Avec Divers Sciences" (1905); "J. de Math." (4) 9 (1893).

Picciati, "Il Politecnico," 41 (1893).

Riquier, "Sur les Systemes d'Equations aux derievees partielles,"(1910).

Severini, "Rend. lst. Lombard." (2) 31 (1898).

Volterra, "Giorn. mat." (1) 19 (1881).

Weierstrass, "Mathematical Werke," Vol. I (1842).

Whittaker and Watson, "Modern Analysis".

BIOGRAPHY

Harris J. Dark, the author of this thesis was born in Maury County, Tennessee, February 8, 1905. He finished high school at Columbia, Tennessee. He did two year's of under-graduate work at David Lipscomb Junior College, one at Vanderbilt University, and one at Randolph-Macon College. From the last of these he received the Bachelor of Arts degree in 1934. During the 1935-36 school year he completed his residence work for the degree of Master of Arts at the University of Richmond, taking Mathematics for his major field and Economics for his minor. During the current school year he has taken an independent study course in Differential Equations at the Universitty of Louisville and written this thesis. He has also done special work in the field of religion, spending one year at the School of Religion of Vanderbilt University and one year at the Union Theological Seminary of Richmond, Virginia.

Dark began preaching at the age of 18 and has been continually employed as a minister since he was 21 years of age, having served the churches of Christ at the following places: Hartsville, Tennessee; Highland Park, Richmond, Virginia; Grant and Summit Streets, Portsmouth, Ohio; and Bardstown Road and Boulevard Napoleon, Louisville, Kentucky, where he is now engaged. He has also held revival meetings in a great many states, including Tennessee, Alabama, Florida, New Jersey, Ohio, Pennsylvania, West Virginia, and Virginia.