

IMPLICIT FUNCTIONS AND THEIR DIFFERENTIALS IN GENERAL ANALYSIS*

BY

T. H. HILDEBRANDT

AND

LAWRENCE M. GRAVES†

Introduction. Implicit function theorems occur in analysis in many different forms, and have a fundamental importance. Besides the classical theorems as given for example in Goursat's *Cours d'Analyse* and Bliss's *Princeton Colloquium Lectures*, and the classical theorems on linear integral equations, implicit function theorems in the domain of infinitely many variables have been developed by Volterra, Evans, Lévy, W. L. Hart and others.‡

The existence and imbedding theorems for solutions of differential equations, as treated for example by Bliss,§ have also received extensions to domains of infinitely many variables by Moulton, Hart, Barnett, Bliss and others.¶ Special properties in the case of linear differential equations in an infinitude of variables have been treated by Hart and Hildebrandt.|| On the other hand, Hahn and Carathéodory** have made important generalizations of the notion of differential equation by removing continuity restrictions on the derivatives and by writing the equations in the form of integral equations.

*Presented to the Society March 25, 1921, and December 29, 1924; received by the editors January 28, 1926.

†National Research Fellow in Mathematics.

‡See Volterra, *Fonctions de Lignes*, Chapter 4; Evans, *Cambridge Colloquium Lectures*, pp. 52-72; Lévy, *Bulletin de la Société Mathématique de France*, vol. 48 (1920), p. 13, Hart, these *Transactions*, vol. 18 (1917), pp. 125-150, where additional references may be found; also vol. 23 (1922), pp. 45-50, and *Annals of Mathematics*, (2), vol. 24 (1922) pp. 29 and 35.

§ *Princeton Colloquium*, and *Bulletin of the American Mathematical Society*, vol. 25 (1918), p. 15.

¶ Moulton, *Proceedings of the National Academy of Sciences*, vol. 1 (1915), p. 350; Hart, papers cited above; Barnett, *American Journal of Mathematics*, vol. 44 (1922), p. 172; Bliss, these *Transactions*, vol. 21 (1920), p. 79.

||Hart, *American Journal of Mathematics*, vol. 39 (1917), p. 407; Hildebrandt, these *Transactions*, vol. 18 (1917), p. 73, and vol. 19 (1918), p. 97.

** Hahn, *Monatshefte für Mathematik und Physik*, vol. 14 (1903), p. 326; Carathéodory, *Vorlesungen über Reelle Funktionen*, p. 666.

These various theories suggest the desirability of a general unifying theory. Lamson began such a general theory in his paper entitled *A general implicit function theorem*.^{*} However, his theory is limited in power, because it contains no theorems on the differentiability of the solutions. Most of the papers on special cases by the other writers mentioned above have a like defect, in that they give, if any, a very incomplete theory of differentiability properties.

The authors propose in the present paper to give a simple general theory of implicit functions and their differentiability, and in a second paper to develop a few special cases, including some new results which find application in the calculus of variations.

The classical implicit function theorems deal with the solution of equations $G(x, y) = 0$, where x and y are points of ordinary space of one or more dimensions. In this paper the authors generalize the theory to the case where x and y are points of abstract spaces of the type discussed by Fréchet. However, no previous knowledge of Fréchet's theory is needed or used. Also the notation is so devised that the reader may readily compare the theory with the classical theory of implicit functions, and interpret it in the finite domain if he so desires.

In the Part I the fundamental postulates and definitions are set down, and certain fundamental propositions are proved. Part II contains the preliminary theorems on the solution of equations in the form $y = F(x, y)$. In Part III, the differential calculus of functions in our abstract spaces is developed. This depends on the notion of total differential, as defined by Stolz[†] for the case of an ordinary function of n variables, applied by Fréchet[‡] in the theory of functionals, and finally developed by Fréchet (independently of the authors of this paper) for the general case.[§] Part IV contains lemmas concerning reciprocal linear functions, and Part V contains the final theorems on the existence and differentiability of implicit functions defined by equations of the form

$$G(x, y) = y_*.$$

^{*} American Journal of Mathematics, vol. 42 (1920), p. 243.

[†] *Grundzüge der Differential- und Integralrechnung*, 1893, vol. 1, pp. 130 ff., 155 ff. Cf. also W. H. Young, Proceedings of the London Mathematical Society, vol. 7 (1909), p. 157; and Fréchet, *Sur la notion de différentielle totale*, *Comptes Rendus du Congrès des Sociétés Savantes en 1914, Sciences*.

[‡] These Transactions, vol. 15 (1914), p. 140.

[§] *Comptes Rendus*, vol. 180 (1925), p. 806, and *Annales de l'École Normale Supérieure*, vol. 41 (1925), p. 293.

I. POSTULATES, DEFINITIONS, AND FUNDAMENTAL PROPOSITIONS

1. **Postulates.** We shall be dealing in the following sections with systems $(\mathfrak{A}, \mathfrak{X}, \| \cdot \|, \oplus, \odot)$ having some or all of the following properties.*

1.11. \mathfrak{A} is either the real number system or the complex number system.

1.12. \mathfrak{X} is a class having at least two distinct elements.

1.13. $\| \cdot \|$ is a function on $\mathfrak{X}\mathfrak{X}$ to the real non-negative part of \mathfrak{A} , i. e., to every pair x_1, x_2 of elements of \mathfrak{X} corresponds a unique positive or zero number denoted by $\|x_1, x_2\|$.

1.14. $\|x_1, x_2\| = \|x_2, x_1\|$ for every x_1, x_2 .

1.15. $\|x_1, x_2\| = 0$ is equivalent to $x_1 = x_2$.

1.16. $\|x_1, x_2\| \leq \|x_1, x_3\| + \|x_3, x_2\|$ for every x_1, x_2, x_3 .

1.2. For every sequence $\{x_n\}$ such that

$$\lim_{\substack{m \rightarrow \infty \\ n \rightarrow \infty}} \|x_m, x_n\| = 0$$

there exists an element x such that

$$\lim_{n \rightarrow \infty} \|x_n, x\| = 0.$$

1.31. \oplus is a function on $\mathfrak{X}\mathfrak{X}$ to \mathfrak{X} , i. e., to every pair x_1, x_2 there corresponds a unique element of \mathfrak{X} denoted by $x_1 \oplus x_2$ and called the sum of x_1 and x_2 .

1.32. \oplus is commutative, i. e., $x_1 \oplus x_2 = x_2 \oplus x_1$ for every x_1, x_2 .

1.33. \oplus is associative, i. e., $(x_1 \oplus x_2) \oplus x_3 = x_1 \oplus (x_2 \oplus x_3)$ for every x_1, x_2, x_3 .

1.34. \odot is a function on $\mathfrak{X}\mathfrak{A}$ to \mathfrak{X} , i. e., to every element x and number a there corresponds a unique element of \mathfrak{X} , denoted by $x \odot a$, and called the product of x by a .

1.35. \odot is associative, i. e., $(x \odot a_1) \odot a_2 = x \odot (a_1 a_2)$ for every x, a_1, a_2 .

1.36. \odot is doubly distributive, i. e., $(x_1 \oplus x_2) \odot a = x_1 \odot a \oplus x_2 \odot a$, and $x \odot (a_1 + a_2) = x \odot a_1 \oplus x \odot a_2$, for every x_1, x_2, x, a, a_1, a_2 .

1.37. $x \odot 1 = x$ for every x .

* For similar sets of postulates, cf. Banach, *Fundamenta Mathematicae*, vol. 3 (1922), p. 133; Hahn, *Monatshefte für Mathematik und Physik*, vol. 32 (1922), p. 3; Fréchet, *Comptes Rendus*, vol. 180 (1925), p. 419. The sets of Banach and of Hahn are somewhat redundant. Our thanks are due to M. H. Ingraham for suggestions tending to eliminate some superfluous postulates.

Notation. The element $x \odot (-1)$ will be denoted by $-x$, and the element $x_1 \oplus (-x_2)$ will be denoted by $x_1 - x_2$. The number $\|x, x \odot 0\|$ will be denoted by $\|x\|$. When no ambiguity can arise, the sign \oplus will be replaced by the ordinary $+$, and \odot will be omitted altogether.

1.38. $\|x_1, x_2\| = \|x_1 - x_2\|$ for every x_1, x_2 .

1.39. $\|xa\| = \|x\| |a|$ for every x and a , where $|a|$ denotes the ordinary absolute value of a .

If a system $(\mathfrak{A}, \mathfrak{X}, \|\cdot\|, \oplus, \odot)$ satisfies the group of postulates 1.11 to 1.16, we shall call \mathfrak{X} a *metric space*; if 1.11 to 1.16 and 1.31 to 1.39 are satisfied we shall call \mathfrak{X} a *linear metric space*; and in either case if 1.2 is also satisfied we shall say that \mathfrak{X} is *complete*.* The elements of a metric space will be called *points*. If we denote the real number system by \mathfrak{R} and the complex number system by \mathfrak{C} , we readily verify that the systems

$$\mathfrak{A} = \mathfrak{R}, \mathfrak{X} = \mathfrak{R}, \|r_1, r_2\| = |r_1 - r_2|, r_1 \oplus r_2 = r_1 + r_2, r_1 \odot r = r_1 r,$$

and

$$\mathfrak{A} = \mathfrak{C}, \mathfrak{X} = \mathfrak{C}, \|c_1, c_2\| = |c_1 - c_2|, c_1 \oplus c_2 = c_1 + c_2, c_1 \odot c = c_1 c,$$

satisfy all the postulates 1.11 to 1.39. Moreover, if \mathfrak{X} is linear metric with \mathfrak{C} as its associated number system, it is also linear metric with \mathfrak{R} as its associated number system.

2. Properties of spaces. A metric space \mathfrak{X} has the following additional properties.

2.11. If the sequence $\{x_n\}$, the points x and x' , and the number a are such that

$$\lim_{n=\infty} \|x_n, x\| = 0,$$

and $\|x_n, x'\| \leq a$ for every n , then $\|x, x'\| \leq a$.

2.12. If a sequence $\{x_n\}$ has a limit, it has only one, i. e., if $\{x_n\}$, x_1 and x_2 are such that

$$\lim_{n=\infty} \|x_n, x_1\| = 0, \quad \lim_{n=\infty} \|x_n, x_2\| = 0,$$

then $x_1 = x_2$.

The first of these follows from 1.14 and 1.16, and the second from 1.14, 1.16 and 1.15.

For a complete metric space \mathfrak{X} we have the following proposition:

*A metric space is one of the classes denoted by (\mathfrak{E}) in Fréchet's thesis, and by (\mathfrak{D}) in his later work. A complete linear metric space is called by Fréchet a space of Banach, or "espace (\mathfrak{D}) vectoriel complet", and is a special case of his "espace affine" for which the postulates are set down in the note cited above.

2.21. If the sequence $\{x_n\}$ is such that the series

$$\sum_n \|x_n, x_{n+1}\|$$

is convergent, then the sequence has a unique limit.

This follows from 1.14, 1.16, 1.2, and 2.12.

A linear metric space \mathfrak{X} has the following additional properties.

2.31. $x_1 \odot 0 = x_2 \odot 0$ for every x_1, x_2 .

Notation. The point $x \odot 0$, which is independent of x by 2.31, will be denoted by x_* .

2.32. $x + x_* = x$ for every x , and x_* is the only point of \mathfrak{X} having this property.

2.33. $xa = x_*$ is equivalent to $a = 0$ or $x = x_*$.

2.34. For every pair x_1, x_2 , there exists uniquely a point x_3 such that $x_1 + x_3 = x_2$, viz., $x_3 = x_2 - x_1$.

2.35. If $a \neq 0$ and $x_1 a = x_2 a$, then $x_1 = x_2$.

2.36. If $x \neq x_*$ and $xa_1 = xa_2$, then $a_1 = a_2$.

2.37. $|\|x_1\| - \|x_2\|| \leq \|x_1 \pm x_2\| \leq \|x_1\| + \|x_2\|$ for every x_1, x_2 .

The property 2.31 follows from 1.11, 1.34, 1.38, 1.35, 1.36, 1.39, and 1.15; 2.32 from 1.11, 2.31, 1.37, 1.36, 1.32; 2.33 from 1.11, 2.31, 1.34, 1.35, 1.37; 2.34 from 1.11, 1.31, 1.32, 1.33, 1.37, 1.36, 2.31, 2.32; 2.35 from 1.11, 1.34, 1.35, 1.37; 2.36 from 1.11, 1.31, 1.36, 2.31, 2.33; 2.37 from 1.11, 1.38, 2.31, 1.14, 1.16, 1.31, 1.32, 1.35, 1.36, 1.37, 1.39, 2.32, 2.34.

3. **Composition of classes.** For definiteness consider two classes \mathfrak{X} and \mathfrak{Y} . Then the composite class $\mathfrak{X}\mathfrak{Y} = (\mathfrak{X}, \mathfrak{Y})$ is defined to be the class of all pairs (x, y) of elements, one from \mathfrak{X} and one from \mathfrak{Y} . If \mathfrak{X} and \mathfrak{Y} are metric spaces, and if we define

$$(3.1) \quad \|w_1, w_2\| \equiv \|(x_1, y_1), (x_2, y_2)\| \equiv \text{greater of } \begin{cases} a\|x_1, x_2\| \\ b\|y_1, y_2\| \end{cases}$$

where a and b are fixed positive constants, then $\mathfrak{X}\mathfrak{Y}$ is also a metric space.* Unless otherwise specified in the sequel, we take $a = b = 1$. If \mathfrak{X} and \mathfrak{Y} are complete metric spaces, so is $\mathfrak{X}\mathfrak{Y}$. If \mathfrak{X} and \mathfrak{Y} are linear metric spaces with the same associated number system \mathfrak{A} , and if we define $w_1 + w_2 \equiv (x_1, y_1) + (x_2, y_2) \equiv (x_1 + x_2, y_1 + y_2)$, $wa \equiv (x, y)a \equiv (xa, ya)$, then $\mathfrak{X}\mathfrak{Y}$ is also a linear metric space. We note that any number of classes may be composed in this way, and that composition may be regarded as an associative process.

*Of course many other definitions might be used for $\|w_1, w_2\|$, e. g., $\|w_1, w_2\| = (\|x_1, x_2\|^p + \|y_1, y_2\|^p)^{1/p}$ ($p > 1$). The one given in the text seems to be the most convenient for our purposes.

4. **Neighborhoods.** For a given point x_0 of a metric space \mathfrak{X} , and a positive number a , we define the neighborhood $(x_0)_a$ to be the set of all points x such that $\|x, x_0\| < a$. When we speak of "a neighborhood" of a point x , we shall always mean a neighborhood of this type. A neighborhood of x may contain no points distinct from x , unless \mathfrak{X} is linear.

5. **Regions.** A region of a metric space \mathfrak{X} is a set \mathfrak{X}_0 of points of \mathfrak{X} such that every point x of \mathfrak{X}_0 has a neighborhood consisting wholly of points of \mathfrak{X}_0 . Regions will be consistently designated by attaching subscripts to the German capitals representing the spaces to which they belong. A neighborhood of a point x is an example of a region. In a composite space $(\mathfrak{X}, \mathfrak{Y}) \equiv \mathfrak{B}$, a neighborhood $(w_0)_c \equiv ((x_0, y_0))_c$ consists of all pairs (x, y) for which x is in $(x_0)_d$ and y is in $(y_0)_e$, where $d = c/a$, $e = c/b$, and a and b are the constants of the definition (3.1) above. The composite $\mathfrak{B}_0 \equiv (\mathfrak{X}_0, \mathfrak{Y}_0)$ of two regions \mathfrak{X}_0 and \mathfrak{Y}_0 is also a region in the composite space $\mathfrak{B} \equiv (\mathfrak{X}, \mathfrak{Y})$, but the converse need not be true. However, the set \mathfrak{Y}_0 of points of \mathfrak{Y} such that (x_0, y) belongs to a region \mathfrak{B}_0 of \mathfrak{B} constitutes a region for each fixed x_0 of \mathfrak{X} , unless the set is empty.

6. **Continuous functions. Relative uniformity.** We shall have frequent use for the notion of relative uniformity, due to E. H. Moore.* Let \mathfrak{X} and \mathfrak{Y} be two metric spaces, and let \mathfrak{P} be a general range. Let F be a function on $\mathfrak{X}_0 \mathfrak{P}$ to \mathfrak{Y} , i. e., F makes correspond to each point x of the region \mathfrak{X}_0 of \mathfrak{X} and each element p of the class \mathfrak{P} one and only one element $F(x, p)$ of \mathfrak{Y} . Let σ be a function on \mathfrak{P} to \mathfrak{A} . Then we say that F is continuous at a point x_0 of \mathfrak{X}_0 uniformly on \mathfrak{P} relative to σ , more briefly, *uniformly* $(\mathfrak{P}; \sigma)$, in case for every $\epsilon > 0$ there exists a $d > 0$ such that, for every x in $(x_0)_d$ and every p of \mathfrak{P} , we have $\|F(x, p), F(x_0, p)\| \leq \epsilon |\sigma(p)|$. We say that F is continuous on \mathfrak{X}_0 uniformly $(\mathfrak{P}; \sigma)$ in case F is continuous at each point of \mathfrak{X}_0 uniformly $(\mathfrak{P}; \sigma)$. Obviously we obtain the definition of ordinary continuity as a special case by taking a singular range \mathfrak{P} and $\sigma(p) = 1$. We say that F is continuous on \mathfrak{X}_0 uniformly $(\mathfrak{X}_0 \mathfrak{P}; \sigma)$ in case for every $\epsilon > 0$ there exists a $d > 0$ such that, for every x_1 and x_2 of \mathfrak{X}_0 satisfying $\|x_1, x_2\| \leq d$ and for every p of \mathfrak{P} we have $\|F(x_1, p), F(x_2, p)\| \leq \epsilon |\sigma(p)|$.

Note that it is important to specify the range of uniformity. In the applications we shall make of the notion of relative uniformity, the range \mathfrak{P} will frequently have as a component a linear metric space \mathfrak{B} , and the scale function σ will then have as a factor the norm $\|w\|$.

*Cf. *Introduction to a Form of General Analysis*, p. 27.

7. **Connected sets.** If \mathfrak{X} is a metric space, we say that a set $\mathfrak{X}^{(0)}$ of points of \mathfrak{X} is a connected set in case for every pair x_1, x_2 of points of $\mathfrak{X}^{(0)}$ there exists a function F on an interval \mathfrak{R}_0 (where \mathfrak{R} is the axis of reals) to $\mathfrak{X}^{(0)}$ which is continuous on \mathfrak{R}_0 and such that $F(r_1) = x_1, F(r_2) = x_2$, where r_1 and r_2 are points of \mathfrak{R}_0 . Not every metric space contains a connected set. However, in a linear metric space, a neighborhood forms an example of a connected region. The composite $(\mathfrak{X}^{(0)}, \mathfrak{Y}^{(0)})$ of two connected sets $\mathfrak{X}^{(0)}$ and $\mathfrak{Y}^{(0)}$ is also a connected set.

8. **Boundary of a set.** We shall say that a point x belongs to the boundary of a set $\mathfrak{X}^{(0)}$ of a metric space \mathfrak{X} in case every neighborhood of x contains both a point of $\mathfrak{X}^{(0)}$ and a point not of $\mathfrak{X}^{(0)}$. The boundary of $\mathfrak{X}^{(0)}$ consists of all such points. By definition of a region, no point of the boundary of a region \mathfrak{X}_0 belongs to \mathfrak{X}_0 . Also if every neighborhood of a point x contains a point of a set $\mathfrak{X}^{(0)}$, then x belongs either to $\mathfrak{X}^{(0)}$ or to the boundary of $\mathfrak{X}^{(0)}$.

II. EXISTENCE AND CONTINUITY OF SOLUTIONS OF EQUATIONS OF THE FORM $y = F(x, y)$

9. Throughout this section we shall denote by \mathfrak{Y} a complete metric space, and by \mathfrak{X} a metric space.

THEOREM 1. *Let the point y_0 of \mathfrak{Y} and the region \mathfrak{X}_0 of \mathfrak{X} and the function F on $\mathfrak{X}_0(y_0)_a$ to \mathfrak{Y} be such that*

(H₁) for every x in \mathfrak{X}_0 there exists a positive constant $k_x < 1$ such that

$$\|F(x, y_1), F(x, y_2)\| \leq k_x \|y_1, y_2\|$$

for every pair y_1, y_2 in $(y_0)_a$;

$$(H_2) \quad \|F(x, y_0), y_0\| < (1 - k_x)a$$

for every x in \mathfrak{X}_0 . Then there exists a unique function Y on \mathfrak{X}_0 to $(y_0)_a$ such that

$$(9.1) \quad Y(x) = F(x, Y(x))$$

for every x in \mathfrak{X}_0 .

This theorem is basic for all the following theorems, and its proof is the only place where a sequence of approximating functions is used. In most applications of this theorem the constant k_x may be taken independent of x . It should be noted also that only an approximate initial solution is required. The proof is as follows.

We define a sequence of approximations by the equations

$$(9.2) \quad Y_1(x) = F(x, y_0), \quad Y_{m+1}(x) = F(x, Y_m(x)) \quad (m > 0).$$

By H_2 we have $\|Y_1(x), y_0\| = (1 - k_x)c_x$, where $c_x < a$, and by H_1 and induction,

$$(9.3) \quad \|Y_{m+1}(x), Y_m(x)\| \leq k_x^m(1 - k_x)c_x \quad (m > 0).$$

From this we obtain

$$(9.4) \quad \|Y_m(x), y_0\| < c_x < a \quad (m > 0),$$

so that the approximations (9.2) are surely defined when x is in \mathfrak{X}_0 .

From the inequality (9.3) and property 2.21 we conclude that there exists a unique function Y on \mathfrak{X}_0 to \mathfrak{Y} such that

$$(9.5) \quad \lim_{m \rightarrow \infty} \|Y_m(x), Y(x)\| = 0.$$

That $Y(x)$ is in $(y_0)_a$ follows from the inequality (9.4) and property 2.11. By H_1 we have

$$\begin{aligned} \|Y, F(x, Y)\| &\leq \|Y, Y_{m+1}\| + \|F(x, Y_m), F(x, Y)\| \\ &\leq \|Y, Y_{m+1}\| + k_x \|Y_m, Y\|, \end{aligned}$$

from which by equation (9.5) and property 1.15 we obtain the desired equation (9.1). To obtain the uniqueness of the solution we again apply H_1 and property 1.15.

We note that the hypotheses of Theorem 1 imply that for each x of the region \mathfrak{X}_0 the function $F(x, y)$ transforms the neighborhood $(y_0)_a$ into a part of itself. For each x the solution $Y(x)$ is the unique invariant point of this transformation. This point of view has been developed by Birkhoff and Kellogg* and by J. L. Holley† for certain special cases.

THEOREM 2. *Let \mathfrak{B}_0 be a region of the composite space $(\mathfrak{X}, \mathfrak{Y})$, and let the function F on \mathfrak{B}_0 to \mathfrak{Y} and the point (x_0, y_0) of \mathfrak{B}_0 be such that*

$$(H_1) \quad y_0 = F(x_0, y_0);$$

(H₂) there exists a positive constant $k < 1$ such that

$$\|F(x, y_1), F(x, y_2)\| \leq k \|y_1, y_2\|$$

for every $(x, y_1), (x, y_2)$ in \mathfrak{B}_0 ;

(H₃) F is continuous in its argument x at (x_0, y_0) . Then the following conclusions hold:

(C₁) for each x there is at most one point (x, y) in \mathfrak{B}_0 which is a solution of the equation.

$$(9.6) \quad y = F(x, y);$$

* These Transactions, vol. 23 (1922), p. 96.

†Harvard thesis, 1924.

(C₂) *there exist a region \mathfrak{X}_1 containing x_0 and a function Y on \mathfrak{X}_1 to \mathfrak{Y} such that the point $(x, Y(x))$ is in \mathfrak{B}_0 and is a solution of (9.6) for every x in \mathfrak{X}_1 ;*

(C₃) *the solution Y is continuous at x_0 ;*

(C₄) *if F is continuous on \mathfrak{B}_0 [uniformly $(\mathfrak{B}_0; 1)$, and if \mathfrak{B}_0 is the composite of two regions \mathfrak{X}_0 and \mathfrak{Y}_0], then the solution Y of equation (9.6) is continuous on its domain of definition [uniformly on that domain].*

C_1 follows as before from H_2 and Postulate 1.15. By the definition of region, there exist positive constants a and b such that the region $((x_0)_b, (y_0)_a)$ is contained in \mathfrak{B}_0 . By H_1 and H_3 , if b is sufficiently small, H_2 of Theorem 1 is satisfied on $(x_0)_b$. Thus C_2 follows from Theorem 1. To obtain C_3 and C_4 (where the parts in brackets form an alternative reading), we use the inequalities

$$\begin{aligned}\|Y_1, Y_2\| &\leq \|F(x_1, Y_1), F(x_1, Y_2)\| + \|F(x_1, Y_2), F(x_2, Y_2)\| \\ &\leq k\|Y_1, Y_2\| + \|F(x_1, Y_2), F(x_2, Y_2)\|, \\ \|Y_1, Y_2\| &\leq (1/(1-k))\|F(x_1, Y_2), F(x_2, Y_2)\|,\end{aligned}$$

where we have set $Y(x_1) = Y_1$, $Y(x_2) = Y_2$. These inequalities are valid in all cases, at least if x_1 is in a sufficiently small neighborhood of x_2 .

Note that C_1 , C_2 , C_4 are still valid if in place of an exact initial solution as assumed in H_1 , we have merely an approximate solution, as in Theorem 1, i. e., if merely $\|F(x_0, y_0), y_0\|$ is sufficiently small.

III. DIFFERENTIALS. THE CLASS $\mathfrak{G}^{(n)}$

Throughout this section and the succeeding ones, we shall assume that we are working with *linear* metric spaces $\mathfrak{X}, \mathfrak{Y}, \mathfrak{Z}, \dots$. In addition we shall frequently add a general range \mathfrak{P} . The associated number system is assumed to be the same for all spaces considered.

10. **Linear functions.** We say that the function F on $\mathfrak{X}\mathfrak{P}$ to \mathfrak{Y} is *distributive* on \mathfrak{X} if for every a_1, a_2, x_1, x_2 and p it is true that

$$F(a_1x_1 + a_2x_2, p) = a_1F(x_1, p) + a_2F(x_2, p)^*.$$

The function F on $\mathfrak{X}\mathfrak{P}$ to \mathfrak{Y} is said to be *modular on \mathfrak{X} uniformly on \mathfrak{P} relative to σ* (or uniformly $(\mathfrak{P}; \sigma)$) in case there exists a constant M such that for every x and p ,

$$\|F(x, p)\| \leq M\|x\| \cdot |\sigma(p)|.$$

*We shall frequently omit the argument p in equations similar to this one especially when p enters "homogeneously."

The minimum of effective values for M , which minimum obviously exists, is called the *modulus* of F .

The function F on $\mathfrak{X}\mathfrak{P}$ to \mathfrak{Y} is said to be *linear* on \mathfrak{X} uniformly $(\mathfrak{P}; \sigma)$ if it is distributive and modular uniformly $(\mathfrak{P}; \sigma)$.

With respect to functions having the modular property we have

LEMMA 10.1. *If F on $\mathfrak{X}_0\mathfrak{Y}\mathfrak{P}$ to \mathfrak{Z} is continuous at x_0 uniformly $(\mathfrak{Y}\mathfrak{P}; \|\gamma\|\sigma)$ and modular on \mathfrak{Y} for $x=x_0$ uniformly $(\mathfrak{P}; \sigma)$, then there exists a neighborhood $(x_0)_a$ of x_0 such that F is modular on \mathfrak{Y} uniformly $((x_0)_a\mathfrak{P}; \sigma)$.*

This is an immediate consequence of the definitions. More generally it is possible to show that if F is continuous on \mathfrak{X}_0 uniformly $(\mathfrak{Y}\mathfrak{P}; \|\gamma\|\sigma)$ and modular on \mathfrak{Y} uniformly $(\mathfrak{P}; \sigma)$ for every x_0 of \mathfrak{X}_0 , then the modulus of F on \mathfrak{Y} is continuous on \mathfrak{X}_0 .

11. The differential. The function F on \mathfrak{X}_0 to \mathfrak{Y} is said to have a differential at x_0 of \mathfrak{X}_0 if there exists a function dF on \mathfrak{X} to \mathfrak{Y} linear on \mathfrak{X} such that the function R on \mathfrak{X}_0 to \mathfrak{Y} defined by the conditions

$$\begin{aligned} F(x_1) - F(x_0) - dF(x_1 - x_0) &= R(x_1)\|x_1 - x_0\| \text{ for } x_1 \neq x_0, \\ \gamma_* &= R(x_1) \text{ for } x_1 = x_0, \end{aligned}$$

is continuous in x_1 at x_0 , i. e.,

$$\lim_{x_1=x_0} \|R(x_1)\| = 0.*$$

It is natural to denote the argument of the function dF by dx . The range of the variable dx is then always the whole space \mathfrak{X} . In the special case $F(x)=x$, we have $dF(dx)=dx$.

We have at once the following result:

LEMMA 11.1. *If F is on \mathfrak{X}_0 to \mathfrak{Y} and has a differential at x_0 of \mathfrak{X}_0 , then this differential is unique and can be obtained as*

$$\lim_{\alpha=0} \frac{F(x_0 + \alpha dx) - F(x_0)}{\alpha} \quad \text{or} \quad \left. \frac{d}{d\alpha} F(x_0 + \alpha dx) \right|_{\alpha=0},$$

the limit being taken in the sense of norm.

This is an immediate consequence of propositions 1.39 and 2.12.

12. The class \mathfrak{C}' . As in ordinary analysis we get a class \mathfrak{C}' of functions F by limiting ourselves to functions F whose differentials dF have certain continuity properties on \mathfrak{X}_0 . We define†

*We follow here the Stolz-Young definition of differential, whose applicability to functional analysis was emphasized by Fréchet. For a discussion of different types of definitions of differentials in the case of functionals cf. Lévy, *Analyse Fonctionnelle*, pp. 50 ff.

†Cf. Bolza, *Variationsrechnung*, p. 13.

F on $\mathfrak{X}_0\mathfrak{P}$ to \mathfrak{Y} is of class \mathfrak{C}' on \mathfrak{X}_0 uniformly $(\mathfrak{P}; \sigma)$ in case there exists a function dF on $\mathfrak{X}_0\mathfrak{X}\mathfrak{P}$ to \mathfrak{Y} having the following properties:

- (1) dF is continuous on \mathfrak{X}_0 uniformly $(\mathfrak{X}\mathfrak{P}; \|dx\|\sigma)$;
- (2) for every x_0 of \mathfrak{X}_0 , dF is linear uniformly $(\mathfrak{P}; \sigma)$;
- (3) for every x_0 of \mathfrak{X}_0 the function R on $\mathfrak{X}_0\mathfrak{X}_0\mathfrak{P}$ to \mathfrak{Y} defined by

$$F(x_1) - F(x_0) - dF(x_0; x_1 - x_0) = R(x_1, x_0)\|x_1 - x_0\| \text{ for } x_1 \neq x_0,$$

$$y_* = R(x_1, x_0) \quad \text{for } x_1 = x_0,$$

is continuous in x_1 at x_0 uniformly $(\mathfrak{P}; \sigma)$.

Obviously this is not the only method of defining the concept "function F of class \mathfrak{C}' ", but it is the simplest direct generalization of the usual definition which adapts itself to elegant results. Another useful concept is obtained by the addition of the range \mathfrak{X}_0 to the range of uniformity in the above conditions on dF . We shall call such a class of functions "class \mathfrak{C}' uniformly $(\mathfrak{X}_0\mathfrak{P}; \sigma)$ ".*

Most of the lemmas which we shall derive hold in two ways, obtained by adding or omitting the class \mathfrak{X}_0 in the range of uniformity in hypothesis and conclusion. In most cases the proof of one lemma thus obtained can be derived from that of the other without difficulty.

LEMMA 12.1. *If F on $\mathfrak{X}_0\mathfrak{P}$ to \mathfrak{Y} is of class \mathfrak{C}' on \mathfrak{X}_0 uniformly $(\mathfrak{P}; \sigma)$ then F is continuous on \mathfrak{X}_0 uniformly $(\mathfrak{P}; \sigma)$. Moreover for every x_0 of \mathfrak{X}_0 there exists a vicinity $(x_0)_a$ and an M such that for all points x of $(x_0)_a$ and all p of \mathfrak{P} it is true that*

$$\|F(x) - F(x_0)\| \leq M\|x - x_0\| \mid \sigma \mid$$

and

$$\|dF(x, dx)\| \leq M\|dx\| \mid \sigma \mid,$$

i. e. F is linear on \mathfrak{X} uniformly $((x_0)_a\mathfrak{P}; \sigma)$.

The first of these is an immediate consequence of the modularity of dF and the application of condition (3) to the inequality

$$\|F(x) - F(x_0)\| \leq \|dF(x_0, x - x_0)\| + \|R(x, x_0)\| \|x - x_0\|.$$

A similar result (without the addition of the uniformity as to σ) follows immediately from the existence of the differential.

The second result is an immediate consequence of Lemma 10.1.

*It is possible to show that if F is of class \mathfrak{C}' uniformly $(\mathfrak{P}; \sigma)$ then the conditions (1) and (2) on dF hold uniformly $(\mathfrak{X}_0\mathfrak{P}; \sigma)$ for every compact subclass \mathfrak{X}_{00} of \mathfrak{X}_0 .

LEMMA 12.2. *If F is of class \mathfrak{C}' on \mathfrak{X}_0 uniformly $(\mathfrak{X}_0\mathfrak{P}; \sigma)$, then F is continuous uniformly $(\mathfrak{X}_0\mathfrak{P}; \sigma)$ and there exist constants a and M such that for every x_1 and x_2 of \mathfrak{X}_0 for which*

$$\|x_1 - x_2\| \leq a$$

and for every p of \mathfrak{P} , it is true that

$$\|F(x_1) - F(x_2)\| \leq M \|x_1 - x_2\| |\sigma|,$$

*i. e., F satisfies a kind of Lipschitz condition on \mathfrak{X}_0 .**

13. **Partial differentials.** If the class \mathfrak{X} is a composite of a finite number of classes $\mathfrak{X}^{(1)}, \mathfrak{X}^{(2)}, \dots, \mathfrak{X}^{(n)}$, and F is on \mathfrak{X}_0 to \mathfrak{Y} , then the existence of the differential of F for $x=x_0$ will imply the existence of partial differentials and we shall have

$$dF(x_0; dx) = \sum_{i=1}^n d_{x^{(i)}} F(x_0; dx^{(i)}),$$

where $dx = (dx^{(1)}, \dots, dx^{(n)})$. In the same way if F is of class \mathfrak{C}' on \mathfrak{X}_0 then the partial differentials have certain continuity properties. We state them in the following lemma for the case in which $n=2$.

LEMMA 13.1. *If $\mathfrak{X}_0 = \mathfrak{X}'_0 \mathfrak{X}''_0$ and F on $\mathfrak{X}_0\mathfrak{P}$ to \mathfrak{Y} is of class \mathfrak{C}' on \mathfrak{X}_0 uniformly $(\mathfrak{P}; \sigma)$ then the partial differentials $d_{x'} F$ and $d_{x''} F$ have the following properties:*

- (0) $dF(x; dx) = d_{x'} F(x; dx') + d_{x''} F(x; dx'')$;
- (1) $d_{x'} F$ and $d_{x''} F$ are continuous on \mathfrak{X}_0 uniformly $(\mathfrak{X}'\mathfrak{P}; \|dx'\|\sigma)$ and $(\mathfrak{X}''\mathfrak{P}; \|dx''\|\sigma)$ respectively;
- (2) $d_{x'} F$ and $d_{x''} F$ are linear in dx' and dx'' respectively, uniformly $(\mathfrak{P}; \sigma)$;
- (3) if $R_{x'}$ on $\mathfrak{X}_0\mathfrak{X}_0\mathfrak{P}$ to \mathfrak{Y} is defined by

$$F(x'_1 x''_1) - F(x'_0 x''_0) - d_{x'} F(x'_0 x''_0; x'_1 - x'_0) = R_{x'}(x_1 x_0) \|x_1 - x_0\|$$

for $x_1 \neq x_0$,

$$y_* = R_{x'}(x_1 x_0) \text{ for } x_1 = x_0,$$

* A similar result for x_1 and x_2 in a neighborhood $(x_0)_a$ of x_0 (a depending on x_0) could be deduced under the hypothesis of uniformity on \mathfrak{P} only, if condition (3) on the class \mathfrak{C}' were replaced by the condition

$$\lim_{x_1=x_0, x_2=x_0} \|R(x_1, x_2)\| = 0 \text{ uniformly } (\mathfrak{P}; \sigma).$$

This extended result is however also deducible from the original condition (3) by an application of Taylor's theorem. Cf. L. M. Graves, *Riemann integration and Taylor's theorem in general analysis*, in the present number of these Transactions.

then $R_{x'}(x_1, x_0)$ considered as a function of x_1 is continuous at x_0 uniformly $(\mathfrak{P}; \sigma)$. A similar condition holds for $R_{x''}(x_1, x_0)$ similarly defined. Conversely, if $d_{x'}F$ and $d_{x''}F$ satisfy the conditions (1), (2), (3) on \mathfrak{X}_0 uniformly $(\mathfrak{P}; \sigma)$ then F is of class \mathfrak{C}' on \mathfrak{X}_0 uniformly $(\mathfrak{P}; \sigma)$ and dF can be defined by condition (0).

The properties (0), (1), and (2) are obvious. For (3) we use the existence of $d_{x'}F$ and $dF = d_{x'}F + d_{x''}F$ uniformly $(\mathfrak{P}; \sigma)$. The proof of the converse is obvious.

In case F is of class \mathfrak{C}' uniformly $(\mathfrak{X}_0\mathfrak{P}; \sigma)$ we have the following simpler lemma:

LEMMA 13.2. *If $\mathfrak{X}_0 = \mathfrak{X}'_0 \mathfrak{X}''_0$ and F is on $\mathfrak{X}_0\mathfrak{P}$ to \mathfrak{Y} , then F is of class \mathfrak{C}' on \mathfrak{X}_0 uniformly $(\mathfrak{X}_0\mathfrak{P}; \sigma)$ if and only if the partial differentials $d_{x'}F$ and $d_{x''}F$ have properties similar to those of dF uniformly $(\mathfrak{X}_0\mathfrak{P}; \sigma)$.*

These lemmas emphasize the fact that the partial differentials are functions on $\mathfrak{X}_0\mathfrak{X}'$ and $\mathfrak{X}_0\mathfrak{X}''$ respectively, and that the limit involved in the total differential is essentially a two-dimensional one.

14. Higher differentials. The class $\mathfrak{C}^{(n)}$. If for some neighborhood of x_0 , $dF(x; d_1x)$ exists and if this function has for every d_1x a differential at x_0 , then F is said to have a second differential, $d^2F(x_0; d_1x, d_2x)$. This function will be distributive in d_1x but not necessarily modular. If the function $d^2F(x_0; d_1x, d_2x)$ is modular in d_2x uniformly $(\mathfrak{X}; \|d_1x\|)$ then obviously d^2F will be bilinear in d_1x and d_2x ; i. e.

$$\begin{aligned} d^2F(x_0; d'_1x + d''_1x, d'_2x + d''_2x) &= d^2F(x_0; d'_1x, d'_2x) \\ &+ d^2F(x_0; d'_1x, d''_2x) + d^2F(x_0; d''_1x, d'_2x) + d^2F(x_0; d''_1x, d''_2x); \end{aligned}$$

and there exists a constant M such that, for every d_1x and d_2x of \mathfrak{X} ,

$$\|d^2F(x_0; d_1x, d_2x)\| \leq M\|d_1x\| \cdot \|d_2x\|.$$

Without further assumptions on d^2F it does not seem possible to prove that d^2F is symmetric in d_1x and d_2x .

Obviously we can extend these definitions to the n th differential and we have, for every $k < n$,

$$d^{n-k}d^kF = d^nF,$$

i. e. the differential operation is associative. But the successive differential operations are not necessarily commutative.

We define the term " F is of class $\mathfrak{G}^{(n)}$ uniformly $([\mathfrak{X}_0]\mathfrak{P}; \sigma)^*$ " by the following recurrence relation:

F is of class $\mathfrak{G}^{(n)}$ uniformly $([\mathfrak{X}_0]\mathfrak{P}; \sigma)$ if F is of class \mathfrak{G}' uniformly $([\mathfrak{X}_0]\mathfrak{P}; \sigma)$ and dF is of class $\mathfrak{G}^{(n-1)}$ uniformly $([\mathfrak{X}_0]\mathfrak{X}\mathfrak{P}; \|d_1x\|\sigma)$.

It is possible to state equivalent definitions as indicated by the following

LEMMA 14.1. *If k is an integer between 0 and n , and if F is of class $\mathfrak{G}^{(k)}$ uniformly $([\mathfrak{X}_0]\mathfrak{P}; \sigma)$ and d^kF is of class $\mathfrak{G}^{(n-k)}$ uniformly $([\mathfrak{X}_0]\mathfrak{X} \cdots \mathfrak{X}; \prod_{i=1}^k \|d_ix\|\sigma)$, then F is of class $\mathfrak{G}^{(n)}$ uniformly $([\mathfrak{X}_0]\mathfrak{P}; \sigma)$ and conversely.*

For convenience we shall omit the range $[\mathfrak{X}_0]\mathfrak{P}$ and the scale σ of uniformity from the discussion. We proceed by induction and assume that we have proved the lemma (necessary and sufficient conditions) for all values of $k \leq m$ and for all $n > m$. We show then

(a) if F is of class $\mathfrak{G}^{(m+1)}$ and $d^{m+1}F$ is of class $\mathfrak{G}^{(n-(m+1))}$ uniformly $(\mathfrak{X} \cdots \mathfrak{X}; \prod_{i=1}^{m+1} \|d_ix\|)$, then F is of class $\mathfrak{G}^{(n)}$.

For if F is of class $\mathfrak{G}^{(m+1)}$ then F is of class \mathfrak{G}' and dF is of class $\mathfrak{G}^{(m)}$ uniformly $(\mathfrak{X}; \|d_1x\|)$. On the other hand, the statement that $d^{m+1}F$ is of class $\mathfrak{G}^{(n-m-1)}$ uniformly $(\mathfrak{X} \cdots \mathfrak{X}; \prod_{i=1}^{m+1} \|d_ix\|)$ is equivalent to $d^m(dF)$ is of class $\mathfrak{G}^{(n-m-1)}$ in the same way. Applying the lemma for $k = m$, it follows that dF is of class $\mathfrak{G}^{(n-m-1+m)}$ or $\mathfrak{G}^{(n-1)}$ uniformly $(\mathfrak{X}; \|d_1x\|)$. This together with the fact that F is of class \mathfrak{G}' gives the result that F is of class $\mathfrak{G}^{(n)}$.

On the other hand

(b) if F is of class $\mathfrak{G}^{(m)}$ and d^mF is of class $\mathfrak{G}^{(n-m)}$ uniformly $(\mathfrak{X} \cdots \mathfrak{X}; \prod_{i=1}^m \|d_ix\|)$, then F is of class $\mathfrak{G}^{(m+1)}$ and $d^{m+1}F$ is of class $\mathfrak{G}^{(n-m-1)}$ uniformly $(\mathfrak{X} \cdots \mathfrak{X}; \prod_{j=1}^{m+1} \|d_jx\|)$, i. e. we can go from $k = m$ to $k = m + 1$.

For from the second condition it follows that d^mF is of class \mathfrak{G}' uniformly $(\mathfrak{X} \cdots \mathfrak{X}; \prod_{i=1}^m \|d_ix\|)$ and $d^{m+1}F$ is of class $\mathfrak{G}^{(n-m-1)}$ uniformly $(\mathfrak{X} \cdots \mathfrak{X}; \prod_{i=1}^{m+1} \|d_ix\|)$. By applying the lemma for $k = m$ and $n = m + 1$, we find that if F is of class $\mathfrak{G}^{(m)}$ and d^mF is of class \mathfrak{G}' uniformly $(\mathfrak{X} \cdots \mathfrak{X}; \prod_{i=1}^m \|d_ix\|)$ then F is of class $\mathfrak{G}^{(m+1)}$. We thus get the two conditions of the conclusion from the two conditions of the hypothesis of our statement (b). This completes the proof of the lemma.

*By enclosing \mathfrak{X}_0 in the bracket [] we shall indicate the possibility of its omission from the range of uniformity.

15. **Functions of functions.** LEMMA 15.1 *If G on $\mathfrak{X}_0\mathfrak{Y}\mathfrak{Z}$ to \mathfrak{Z} is of class \mathfrak{E}' on \mathfrak{X}_0 uniformly ($\mathfrak{Y}\mathfrak{Z}; \|y\| \sigma$) and linear on \mathfrak{Y} for every x of \mathfrak{X}_0 uniformly ($\mathfrak{Z}; \sigma$); if moreover H on \mathfrak{U}_0 to \mathfrak{X}_0 is of class \mathfrak{E}' on \mathfrak{U}_0 , and K on $\mathfrak{U}_0\mathfrak{V}$ to \mathfrak{Y} is of class \mathfrak{E}' on \mathfrak{U}_0 uniformly ($\mathfrak{V}; \|v\|$) and linear on \mathfrak{V} for each u of \mathfrak{U}_0 , then*

$$G(u, v, p) = G(H(u), K(u, v), p)$$

on $\mathfrak{U}_0\mathfrak{V}\mathfrak{Z}$ to \mathfrak{Z} is of class \mathfrak{E}' on \mathfrak{U}_0 uniformly ($\mathfrak{V}\mathfrak{Z}; \|v\| \sigma$) and linear on \mathfrak{V} for each u of \mathfrak{U}_0 uniformly ($\mathfrak{Z}; \sigma$); moreover

$$d_u G(u, v, p; du) = d_x G(H(u), K(u, v), p; dH(u; du)) + G(H(u), dK(u, v; du), p).$$

The fact that $G(H(u), K(u, v), p)$ is linear on \mathfrak{V} for each u of \mathfrak{U}_0 uniformly ($\mathfrak{Z}; \sigma$) is obvious. We show that G on $\mathfrak{U}_0\mathfrak{V}\mathfrak{Z}$ to \mathfrak{Z} is of class \mathfrak{E}' uniformly ($\mathfrak{V}\mathfrak{Z}; \|v\| \sigma$) by showing that $d_u G$ satisfies the conditions imposed in the definition of the class \mathfrak{E}' . For simplicity we shall omit the range \mathfrak{Z} and the scale function σ from the argument. It is easy to see that their addition does not change the form of the reasoning.

We shall show in the first place that $d_u G$ is continuous on \mathfrak{U}_0 uniformly ($\mathfrak{U}\mathfrak{V}; \|du\| \cdot \|v\|$). We consider for this purpose each term of $d_u G$ separately and study the difference

$$\begin{aligned} & dG(H(u), K(u, v); dH(u; du)) - dG(H(u_0), K(u_0, v); dH(u_0; du)) \\ &= [dG(H(u), K(u, v); dH(u; du)) - dG(H(u_0), K(u, v); dH(u; du))] \\ &\quad + dG(H(u_0), K(u, v); dH(u; du) - dH(u_0; du)) \\ &\quad + dG(H(u_0), K(u, v) - K(u_0, v); dH(u_0; du)). \end{aligned}$$

In order to show that each of these three terms approaches z_* , we assume that u has been chosen in the vicinity of u_0 on which, by Lemmas 10.1 and 12.1, $dH(u; du)$ is uniformly modular in du , and $K(u, v)$ uniformly modular in v . Then the continuity of dG uniformly relative to

$$\|dH(u; du)\| \cdot \|K(u, v)\|$$

and the continuity of H suffice to make the first difference approach z_* uniformly ($\mathfrak{U}\mathfrak{V}; \|du\| \cdot \|v\|$); the modularity of dG uniformly relative to $\|K(u, v)\|$ and the continuity of $dH(u; du)$ uniformly relative to $\|du\|$ produce the approach to z_* uniformly ($\mathfrak{U}\mathfrak{V}; \|du\| \cdot \|v\|$) for the second term, while the additional continuity of $K(u, v)$ uniformly ($\mathfrak{V}; \|v\|$) produces the same result in the third term.

The continuity of $G(H(u); dK(u, v; du))$ is proved along entirely similar lines.

The modularity of $d_u G$ follows immediately from that of $d_x G$, G , dH and dK .

We consider finally

$$\begin{aligned}
 R(u_1, u_0) \|u_1 - u_0\| &= G(H(u_1), K(u_1, v)) - G(H(u_0), K(u_0, v)) \\
 &\quad - d_x G(H(u_0), K(u_0, v); dH(u_0; u_1 - u_0)) \\
 &\quad - G(H(u_0), dK(u_0, v; u_1 - u_0)) \\
 &= [G(H(u_1), K(u_1, v) - K(u_0, v)) \\
 &\quad - G(H(u_0), dK(u_0, v; u_1 - u_0))] \\
 &\quad + [G(H(u_1), K(u_0, v)) - G(H(u_0), K(u_0, v)) \\
 &\quad - dG(H(u_0), K(u_0, v); dH(u_0; u_1 - u_0))].
 \end{aligned}$$

The first group of terms we rewrite in the form

$$\begin{aligned}
 &[G(H(u_1), K(u_1, v) - K(u_0, v) - dK(u_0, v; u_1 - u_0))] \\
 &\quad + [G(H(u_1), dK(u_0, v; u_1 - u_0)) - G(H(u_0), dK(u_0, v; u_1 - u_0))].
 \end{aligned}$$

If now we assume that u_1 is in such a neighborhood of u_0 that $H(u_1)$ lies in the neighborhood of $H(u_0)$ for which, by Lemma 10.1, G is uniformly modular on \mathfrak{Y} , it follows that the norm of the first expression is less than or equal to

$$M_G \|R_K(u_1, u_0)\| \cdot \|u_1 - u_0\|$$

in which the coefficient of $\|u_1 - u_0\|$ approaches zero uniformly ($\mathfrak{B}; \|v\|$) as u_1 approaches u_0 . For the second expression we utilize the continuity of G uniformly ($\mathfrak{Y}; \|y\|$), i. e. ($\mathfrak{Y}; \|dK(u_0, v; u_1 - u_0)\|$) which in turn can

be replaced by uniformly $(\mathfrak{Y}; M_k \|v\| \cdot \|u - u_0\|)$ so that the second term when divided by $\|u_1 - u_0\|$ will approach z_* uniformly $(\mathfrak{B}; \|v\|)$ as u_1 approaches u_0 .

The second group of terms in the expression for $R(u_1, u_0)\|u_1 - u_0\|$ can be replaced by

$$R_G(H(u_1), H(u_0); K(u_0, v))\|H(u_1) - H(u_0)\| \\ + dG(H(u_0), K(u_0, v); R_H(u_1, u_0)\|u_1 - u_0\|).$$

By Lemma 12.1 if u_1 is in a sufficiently small neighborhood of u_0 there exists an M such that

$$\|H(u_1) - H(u_0)\| \leq M\|u_1 - u_0\|.$$

Hence the first term when divided by $\|u_1 - u_0\|$ will approach z_* uniformly $(V; \|v\|)$. The same result for the second term follows from the modular properties of dG . This completes the proof of

$$\lim_{u_1 \rightarrow u_0} \|R(u_1, u_0)\| = 0 \text{ uniformly } (\mathfrak{B}; \|v\|).$$

We note that the addition of further linearity in the function K will induce corresponding linearity in G , i. e., if $K(u; v_1, \dots, v_n)$ on $\mathfrak{U}_0 \mathfrak{B}_1 \cdots \mathfrak{B}_n$ is of class \mathfrak{E}' on \mathfrak{U}_0 uniformly $(\mathfrak{B}_1 \cdots \mathfrak{B}_n; \prod_{i=1}^n \|v_i\|)$ and linear on each \mathfrak{B}_i uniformly $(\mathfrak{B}_1 \cdots \mathfrak{B}_{i-1} \mathfrak{B}_{i+1} \cdots \mathfrak{B}_n; \prod_{i \neq j} \|v_j\|)$ then G has the same properties as K .

It is possible to extend this lemma still further. To simplify the statement we introduce the following terminology: G on $\mathfrak{X}_0 \mathfrak{Y}_1 \cdots \mathfrak{Y}_m \mathfrak{P}$ has the property $P^{(n)}$ on $\mathfrak{X}_0 \mathfrak{Y}_1 \cdots \mathfrak{Y}_m$ uniformly $(\mathfrak{P}; \sigma)$, in case G is of class $\mathfrak{E}^{(n)}$ on \mathfrak{X}_0 uniformly $(\mathfrak{Y}_1 \cdots \mathfrak{Y}_m \mathfrak{P}; \prod_{i=1}^m \|y_i\| \sigma)$ and for each x of \mathfrak{X}_0 , G is linear on \mathfrak{Y}_1 uniformly $(\mathfrak{Y}_1 \cdots \mathfrak{Y}_{i-1} \mathfrak{Y}_{i+1} \cdots \mathfrak{Y}_m \mathfrak{P}; \prod_{i \neq j} \|y_j\| \sigma)$.

Obviously if G has the property P' and dG has the property $P^{(n-1)}$ uniformly $(\mathfrak{X}; \|dx\|)$ then G has the property $P^{(n)}$.

We then have the following lemma:

LEMMA 15.2. *If G on $\mathfrak{X}_0 \mathfrak{Y}_1 \cdots \mathfrak{Y}_k \mathfrak{P}$ to \mathfrak{Z} has the property $P^{(n)}$ on $\mathfrak{X}_0 \mathfrak{Y}_1 \cdots \mathfrak{Y}_k$ uniformly $(\mathfrak{P}; \sigma)$; if, moreover, H on \mathfrak{U}_0 to \mathfrak{X}_0 is of class $\mathfrak{E}^{(n)}$ on \mathfrak{U}_0 , and K_i on $\mathfrak{U}_0 \mathfrak{B}_{i1} \cdots \mathfrak{B}_{ij_i}$ to \mathfrak{Y}_i has the property $P^{(n)}$ on $\mathfrak{U}_0 \mathfrak{B}_{i1} \cdots \mathfrak{B}_{ij_i}$, then G on $\mathfrak{U}_0 \mathfrak{B}_{11} \cdots \mathfrak{B}_{kij_k} \mathfrak{P}$ has the property $P^{(n)}$ on $\mathfrak{U}_0 \mathfrak{B}_{11} \cdots \mathfrak{B}_{kij_k}$ uniformly $(\mathfrak{P}; \sigma)$.*

The proof for $n=1$ and $k=1$ has been given in Lemma 15.1 and the appended remark. Assume that it has been proved for values $n=1$ and

$k \leq m$ and show that it holds for $k = m + 1$. This is a result of applying the lemma for $k = m$ and $k = 1$ to the following functions:

$$\begin{aligned}\bar{G}(u, y_{m+1}, v_{11}, \dots, v_{mim}, p) &= G(H, K_1, \dots, K_m, y_{m+1}, p), \\ \bar{H}(u) &= u, \\ \bar{K}(u, v_{m+1,1}, \dots, v_{m+1,im+1}) &= y_{m+1} = K_{m+1}(u, v_{m+1,1}, \dots, v_{m+1,im+1}).\end{aligned}$$

To prove the lemma for any n , assume that it has been proved for $n \leq m$ and extend it to the case $n = m + 1$. It is obviously sufficient to show that $d_u G$ has the property $P^{(m)}$ uniformly $(\mathfrak{P}; \sigma)$. Now

$$\begin{aligned}d_u G &= d_x G(H, K_1, \dots, K_k, p; dH(du)) \\ &+ \sum_{i=1}^k G(H, K_1, \dots, dK_i(du), K_k, p).\end{aligned}$$

Obviously H is of class $\mathfrak{C}^{(m)}$, and $d_x G, dH, K_1, \dots, K_k, dK_1, \dots, dK_k$ have the property $P^{(m)}$ in their respective arguments. From the relation of $P^{(m)}$ to the class $\mathfrak{C}^{(m)}$ it follows that $d_u G$ has the property $P^{(m)}$ uniformly $(\mathfrak{U}\mathfrak{P}; \|du\|\sigma)$ and consequently G has the property $P^{(m+1)}$ uniformly $(\mathfrak{P}; \sigma)$.

As a special case of this lemma we note the following

LEMMA 15.3. *If G is on $\mathfrak{X}_0\mathfrak{P}$ to \mathfrak{Z} of class $\mathfrak{C}^{(n)}$ on \mathfrak{X}_0 uniformly $(\mathfrak{P}; \sigma)$ and H is on \mathfrak{U}_0 to \mathfrak{X}_0 of class $\mathfrak{C}^{(n)}$ on \mathfrak{U}_0 , then $G(H(u), p)$ on $\mathfrak{U}_0\mathfrak{P}$ to \mathfrak{Z} is of class $\mathfrak{C}^{(n)}$ on \mathfrak{U}_0 uniformly $(\mathfrak{P}; \sigma)$.*

We need only apply the preceding lemma to the following situation:

$\mathfrak{B} = \mathfrak{Y} = \mathfrak{R}$ = the class of real numbers,

$K(u, r) = r$ for every u of \mathfrak{U}_0 ,

$G(x, r) = rG(x)$ and $H(u) = H(u)$.

We note that if G is of class \mathfrak{C} and H of class \mathfrak{C}' then

$$d_u G(H(u)) = d_x G(H(u); dH(u, du)),$$

a function of the type considered in Lemma 15.1.

In Lemmas 15.1, 15.2 and 15.3 it is possible to get a parallel group of lemmas by inserting the classes \mathfrak{X}_0 and \mathfrak{U}_0 in the range of uniformity. In this case, especially for Lemma 15.1, the proof is simplified in a few places. As a matter of fact it is possible in this case to prove Lemma 15.3 directly by a line of reasoning which considers dG as a function of class $\mathfrak{C}^{(m)}$ of a two-partite argument, each argument in turn being a function on a two-partite class.

IV. ON THE RECIPROCAL OF LINEAR FUNCTIONS

16. In this section we shall assume that \mathfrak{Y} is a complete linear metric space and \mathfrak{X} is a linear metric space.

We say that a function K on \mathfrak{Y} to \mathfrak{Y} , linear on \mathfrak{Y} , has a *reciprocal* if there exists a function L on \mathfrak{Y} to \mathfrak{Y} linear on \mathfrak{Y} satisfying the conditions

$$K(L(y)) = L(K(y)) = y$$

for every y of \mathfrak{Y} .

We note that if such an L exists, then it is unique. For from

$$K(L_1(y)) = K(L_2(y)) = y$$

follows

$$L_2KL_1 = L_2KL_2 \text{ or } L_1(y) = L_2(y).$$

Also from the existence of the reciprocal L it follows that if

$$L(y) = y_*$$

for a given y , then $y = y_*$. Conversely if there exists an L on \mathfrak{Y} to \mathfrak{Y} linear on \mathfrak{Y} satisfying the two conditions

$$L(K(y)) = y \text{ for every } y \text{ of } \mathfrak{Y},$$

$$L(y) = y_* \text{ if and only if } y = y_*,$$

then

$$K(L(y)) = y \text{ for every } y \text{ of } \mathfrak{Y}.$$

For

$$L(KLy - y) = LKLy - Ly = Ly - Ly = y_*.$$

We take up first the reciprocal of a function which is suggested by the linear integral equation of the second kind in the following

LEMMA 16.1. *If G on $\mathfrak{X}_0\mathfrak{Y}$ to \mathfrak{Y} is linear on \mathfrak{Y} for each w of \mathfrak{X}_0 with modulus $M(G, w) < 1$, then for every w of \mathfrak{X}_0 there exists a reciprocal of $y - G(w, y)$. If G is continuous on \mathfrak{X}_0 uniformly ($\mathfrak{Y}; \|y\|$) then the reciprocal is also continuous uniformly ($\mathfrak{Y}; \|y\|$).*

The proof of the first part of this lemma follows the lines of the successive substitution method used in Theorem 1, or the lemma can be shown to be a corollary of Theorem 1, which amounts essentially to the same thing. From either point of view, it appears that if $G_n(w, y) = G(w, G_{n-1}(w, y))$; $G_1(w, y) = G(w, y)$, then for all (w, y) of $\mathfrak{X}_0\mathfrak{Y}$ the reciprocal of $y - G(w, y)$ is

$$y - H(w, y) = y + \sum_1^{\infty} G_n(w, y),$$

and the modulus

$$M(H, w) \leq \frac{M(G, w)}{1 - M(G, w)}$$

Obviously H satisfies the relations

$$H(w, y) + G(w, y) = G(w, H(w, y)) = H(w, G(w, y))$$

which are reminiscent of the reciprocal relations of linear integral equations.

To show that the continuity of G on \mathfrak{X}_0 uniformly $(\mathfrak{Y}; \|y\|)$ implies a similar continuity in the function H we utilize the first of these reciprocal relations to obtain

$$\begin{aligned} H(w_1, y) + G(w_1, y) - G(w_1, H(w_1, y)) \\ = H(w_0, y) + G(w_0, y) - G(w_0, H(w_0, y)) \end{aligned}$$

from which follows

$$\begin{aligned} \|H(w_1, y) - H(w_0, y)\| &\leq \|G(w_1, y) - G(w_0, y)\| + \|G(w_1, H(w_1, y)) \\ &\quad - H(w_0, y)\| + \|G(w_1, H(w_0, y)) - G(w_0, H(w_0, y))\|. \end{aligned}$$

Now from the continuity of $G(w, y)$ on \mathfrak{X}_0 uniformly $(\mathfrak{Y}; \|y\|)$ it follows that there will exist a vicinity $(w_0)_a$ of w_0 such that if w_1 is in this vicinity then $M(G, w_1) < k < 1$, so that for such w_1

$$\begin{aligned} \|H(w_1, y) - H(w_0, y)\|(1 - k) &\leq \|G(w_1, y) - G(w_0, y)\| \\ &\quad + \|G(w_1, H(w_0, y)) - G(w_0, H(w_0, y))\|. \end{aligned}$$

The continuity properties of G applied to this inequality give us corresponding continuity properties of H .

By the same method we can obviously show that if G is continuous on \mathfrak{X}_0 uniformly $(\mathfrak{X}_0\mathfrak{Y}; \|y\|)$ then for every w_0 there exists a vicinity $(w_0)_a$ such that H is continuous on $(w_0)_a$ uniformly $((w_0)_a\mathfrak{Y}; \|y\|)$. This vicinity can be extended to include the region \mathfrak{X}_0 if there exists a $k < 1$ such that for all w of \mathfrak{X}_0

$$M(G, w) < k.$$

LEMMA 16.2. *Suppose that K on $\mathfrak{X}_0\mathfrak{Y}$ to \mathfrak{Y} is linear on \mathfrak{Y} for each w of \mathfrak{X}_0 ; that it has a reciprocal L_0 for $w = w_0$, and is continuous in w at $w = w_0$ uniformly $(\mathfrak{Y}; \|y\|)$; then there exists a constant a such that for each w of the neighborhood $(w_0)_a$, K has a reciprocal, which is continuous at w_0 uniformly $(\mathfrak{Y}; \|y\|)$.*

Consider the function

$$G(w, y) = L_0(K(w_0, y) - K(w, y)) = y - L_0K(w, y).$$

Using the continuity of K at w_0 , we can select a k and an a such that for all w in $(w_0)_a$ we have

$$\|K(w_0, y) - K(w, y)\| < k\|y\|, \quad kM(L_0) < 1.$$

Then

$$M(G, w) < k_1 < 1$$

and so by Lemma 16.1,

$$L_0K(w, y) = y - G(w, y)$$

has on $(w_0)_a$ a reciprocal L_1 continuous at $w=w_0$ uniformly $(\mathfrak{Y}; \|y\|)$. Then for every y of \mathfrak{Y} and w of $(w_0)_a$,

$$L_1L_0Ky = L_0KL_1y = y,$$

where for convenience we have omitted the w . Then L_1L_0 is a left hand reciprocal of K . That it is also a right hand reciprocal follows from

$$L_0KL_1L_0y = L_0y \quad \text{or} \quad L_0(KL_1L_0y - y) = y_*$$

and the property of L_0 as reciprocal of $K(w_0, y)$. The continuity of L_1L_0 at $w=w_0$ uniformly $(\mathfrak{Y}; \|y\|)$ follows from that of L_1 .

If K is also continuous on \mathfrak{X}_0 uniformly $([\mathfrak{X}_0]\mathfrak{Y}; \|y\|)$ then this reciprocal will be continuous on $(w_0)_a$ uniformly $([(w_0)_a]\mathfrak{Y}; \|y\|)$.

Differential properties of K carry over to the reciprocal as we show in

LEMMA 16.3. *If K on $\mathfrak{X}_0\mathfrak{Y}$ to \mathfrak{Y} is linear on \mathfrak{Y} for each w of \mathfrak{X}_0 , has a reciprocal L_0 for $w=w_0$ and is of class $C^{(n)}$ on \mathfrak{X}_0 uniformly $([\mathfrak{X}_0]\mathfrak{Y}; \|y\|)$, then there exists a constant a such that the reciprocal L is of class $C^{(n)}$ on $(w_0)_a$ uniformly $([(w_0)_a]\mathfrak{Y}; \|y\|)$.*

We select the constant a in such a way that the function $L(w, y)$ is uniformly modular on $(w_0)_a$, which is always possible as can be seen from Lemma 10.1. We then prove this lemma first for the case $n=1$ by showing that

$$(16.4) \quad dL(w, y; dw) = -L(w, dK(w, L(w, y); dw))$$

satisfies the three conditions imposed on differentials of functions of the class \mathfrak{C}' . The relation (16.4) can be obtained by applying Lemma 15.1 heuristically to the reciprocal relation

$$K(w, L(w, y)) = y.$$

The continuity of dL can be shown by following the method used, or the result obtained in the proof of the corresponding part of Lemma 15.1. Similarly the modularity of dL is deducible from that of L and dK . It remains to show that dL satisfies condition (3). From the reciprocal relation between K and L it follows that for w_1 and w_2 in $(w_0)_a$ it is true that

$$K(w_1, L(w_1, y)) = K(w_2, L(w_2, y))$$

or

$$\begin{aligned} K(w_1, L(w_2, y)) - K(w_1, L(w_1, y)) &= - (K(w_2, L(w_2, y)) - K(w_1, L(w_2, y))) \\ &= - dK(w_1, L(w_2, y); w_2 - w_1) \\ &\quad - R_K(w_2, w_1, L(w_2, y)) \|w_2 - w_1\|. \end{aligned}$$

Then

$$\begin{aligned} L(w_2, y) - L(w_1, y) &= - L(w_1, dK(w_1, L(w_2, y); w_2 - w_1)) \\ &\quad - L(w_1, R_K(w_2, w_1, L(w_2, y))) \|w_2 - w_1\|, \end{aligned}$$

so that

$$\begin{aligned} R_L(w_2, w_1, y) \|w_2 - w_1\| &= L(w_2, y) - L(w_1, y) - dL(w_1, y; w_2 - w_1) \\ &= - L(w_1, dK(w_1, L(w_2, y); w_2 - w_1)) \\ &\quad - L(w_1, R_K(w_2, w_1, L(w_2, y))) \|w_1 - w_1\|. \end{aligned}$$

Now by Lemma 12.1, K is continuous at w_1 uniformly $(\mathfrak{Y}; \|y\|)$ and so by Lemma 16.2 L is continuous at w_1 uniformly $(\mathfrak{Y}; \|y\|)$. Further, from the modularity of dK and L it follows that

$$\begin{aligned} \|L(w_1, dK(w_1, L(w_2, y) - L(w_1, y); w_2 - w_1))\| \\ \leq M(L)M(dK)\|L(w_2, y) - L(w_1, y)\| \|w_2 - w_1\|, \end{aligned}$$

so that the coefficient of $\|w_2 - w_1\|$ approaches zero uniformly $(\mathfrak{Y}; \|y\|)$ as w_2 approaches w_1 . Similarly, from the uniform modularity of L on $(w_0)_a$ and the manner of approach of R_K to y_* it follows that

$$\|L(w_1, R_K(w_2, w_1, L(w_1, y)))\|$$

approaches zero uniformly $(\mathfrak{Y}; \|y\|)$. This completes the proof that R_L approaches y_* uniformly $(\mathfrak{Y}; \|y\|)$ as w_2 approaches w_1 .

We prove the general case by induction, i. e. assume that the lemma holds for $n \leq m$ and show its validity for $n = m + 1$. Assume then that K is of class $\mathfrak{E}^{(m+1)}$ uniformly $(\mathfrak{Y}; \|y\|)$. Then dK is of class $\mathfrak{E}^{(m)}$ on \mathfrak{X}_0 uniformly $(\mathfrak{Y}\mathfrak{X}; \|y\| \cdot \|dw\|)$ and also L is of class $\mathfrak{E}^{(m)}$ on \mathfrak{X}_0 uniformly $(\mathfrak{Y}; \|y\|)$.

Then by applying Lemma 15.2 we get that $dK(w, L(w, y); dw)$ is of class $\mathfrak{E}^{(m)}$ on \mathfrak{X}_0 uniformly ($\mathfrak{Y}\mathfrak{B}; \|y\| \cdot \|dw\|$) and so by a second application of Lemma 15.2 it follows that $L(w, dK(w, L(w, y); dw)) = -dL(w, y; dw)$ has the same property. This completes the proof of the lemma.

V. EXISTENCE AND DIFFERENTIABILITY OF SOLUTIONS OF
EQUATIONS OF THE FORM $G(x, y) = y_*$

17. Neighborhood theorems. THEOREM 3. Let \mathfrak{X} be a metric space and \mathfrak{Y} a complete linear metric space. Let the point y_0 of \mathfrak{Y} and the region \mathfrak{X}_0 of \mathfrak{X} and the function G on $\mathfrak{X}_0(y_0)_a$ to \mathfrak{Y} be such that

(H_1) there exist a linear function K on \mathfrak{Y} to \mathfrak{Y} with a reciprocal L , and a positive constant M_0 such that, if $M(L)$ denote the modulus of L , then

$$M_0 M(L) < 1, \quad \|K(y_1 - y_2) - G(x, y_1) + G(x, y_2)\| \leq M_0 \|y_1 - y_2\|$$

for every x in \mathfrak{X}_0 and every y_1 and y_2 in $(y_0)_a$;

(H_2) there exists a positive constant $c < a$ such that

$$M(L) \|G(x, y_0)\| \leq (1 - M_0 M(L))c$$

for every x in \mathfrak{X}_0 .

Then for every x in \mathfrak{X}_0 there exists a unique solution $Y(x)$ of the equation

$$(17.1) \quad G(x, y) = y_*.$$

Note that in Theorem 3, only an approximate initial solution, and also only an approximate differential at the approximate initial solution are required. The proof is as follows:

We define a new function F on $\mathfrak{X}_0(y_0)_a$ to \mathfrak{Y} by the equality

$$F(x, y) = y - L(G(x, y)).$$

Then the equation

$$(17.2) \quad y = F(x, y)$$

is equivalent to (17.1). To equation (17.2) we can apply Theorem 1. For by H_1 we have

$$F(x, y_1) - F(x, y_2) = L(K(y_1 - y_2) - G(x, y_1) + G(x, y_2)),$$

$$\|F(x, y_1) - F(x, y_2)\| \leq M(L) M_0 \|y_1 - y_2\|,$$

so that the function F satisfies H_1 of Theorem 1. That the hypothesis H_2 of Theorem 1 also holds for the function F is readily verified.

THEOREM 4. Let \mathfrak{X} be a linear metric space, and \mathfrak{Y} a complete linear metric space, and let \mathfrak{W}_0 be a region of the composite space $\mathfrak{W} = (\mathfrak{X}, \mathfrak{Y})$. Let the function G on \mathfrak{W}_0 to \mathfrak{Y} and the point (x_0, y_0) of \mathfrak{W}_0 satisfy

$$(H_1) \quad G(x_0, y_0) = y_*;$$

$$(H_2) \quad G \text{ is of the class } \mathfrak{E}^{(n)} \text{ on } \mathfrak{W}_0 \text{ [uniformly on } \mathfrak{W}_0];$$

$$(H_3) \quad d_y G(x_0, y_0; dy) \equiv K_0(dy) \text{ has a reciprocal } L_0.$$

Then there exist positive constants a and b and a function Y on $(x_0)_b$ to $(y_0)_a$ with the following properties:

$$(C_1) \quad \text{the region } ((x_0)_b, (y_0)_a) \text{ is contained in } \mathfrak{W}_0;$$

$$(C_2) \quad \text{the point } (x, Y(x)) \text{ is a solution of the equation}$$

$$(17.3) \quad G(x, y) = y_*$$

for every x in $(x_0)_b$, and there is no other solution with the same x , having y in $(y_0)_a$;

$$(C_3) \quad \text{the differential } d_y G(x, Y(x); dy) \text{ has a reciprocal for every } x \text{ in } (x_0)_b;$$

$$(C_4) \quad \text{the function } Y \text{ is of the class } \mathfrak{E}^{(n)} \text{ on } (x_0)_b \text{ [uniformly on } (x_0)_b].$$

The parts in brackets in H_2 and C_4 constitute an alternative reading. We shall prove only the theorem with the parts in brackets omitted. The proof of the other theorem is parallel, except that no recourse to Taylor's theorem is required. We proceed as follows:

Let $M(L_0)$ be the modulus of the reciprocal L_0 of K_0 , and let M_0 be a constant satisfying the condition $0 < M_0 M(L_0) < 1$. By Taylor's theorem* we have, for every $(x, y_1), (x, y_2)$ in a sufficiently restricted neighborhood $((x_0, y_0)_a)$,

$$G(x, y_1) - G(x, y_2) = \int_0^1 d_y G(x, y_2 + (y_1 - y_2)r; y_1 - y_2) dr.$$

If we take a sufficiently small, the inequality

$$\|G(x, y_1) - G(x, y_2) - K_0(y_1 - y_2)\| \leq M_0 \|y_1 - y_2\|$$

follows from the continuity of $d_y G$ at (x_0, y_0) and the properties of Riemann integrals.† Therefore the hypothesis H_1 of Theorem 3 is fulfilled on this neighborhood. By Lemma 12.1 there exists a positive constant $b \leq a$ such that H_2 of Theorem 3 is fulfilled on $(x_0)_b$. Hence the conclusions C_1 and C_2 follow from Theorem 3.

*See Graves, *Riemann integration and Taylor's theorem in general analysis*, already cited.

†Graves, loc. cit.

By Lemma 12.1 and Theorem 2, we find that the solution $Y(x)$ is continuous on $(x_0)_b$. Then by Lemma 16.2, if the constant b is sufficiently restricted, $d_y G(x, Y(x); dy)$ has a reciprocal $L(x; dy)$ which is continuous on $(x_0)_b$ uniformly $(\mathfrak{Y}; \|dy\|)$ and linear on \mathfrak{Y} uniformly on $(x_0)_b$. This gives us C_3 .

In proving C_4 for $n=1$, we show that the differential dY is given by the formula

$$dY(x; dx) = -L(x; d_x G(x, Y(x); dx)).$$

The continuity and linearity of dY follow readily from the corresponding properties of L and $d_x G$. Since G is of class \mathfrak{G}' and $G(x, Y) = y_*$, we have

$$y_* = d_y G(x, Y; Y_1 - Y) + d_x G(x, Y; x_1 - x) + R_G(W_1, W) \|W_1 - W\|.$$

By applying the operation L to both sides of this equation, we obtain

$$(17.4) \quad y_* = Y_1 - Y - dY(x; x_1 - x) + L(x; R_G(W_1, W)) \|W_1 - W\|.$$

Here we have set $W = (x, Y(x))$, etc. If the ratio $r = \|W_1 - W\| \div \|x_1 - x\|$ is bounded when x_1 is in a sufficiently small neighborhood of x , then the remainder $R_Y(x_1, x) = -L(x; R_G(W_1, W))r$ is continuous at $x_1 = x$, since we have already shown that the function L is linear and the function Y is continuous. To show that r is bounded, we evidently need consider only the points x_1 at which $\|W_1 - W\| = \|Y_1 - Y\|$. At these points we have from equation (17.4),

$$\|Y_1 - Y\| \leq M(dY) \|x_1 - x\| + M(L) \|R_G\| \|Y_1 - Y\|.$$

Select a constant c so small that whenever x_1 is in $(x)_c$ we have $M(L) \|R_G\|$ less than a constant $e < 1$. Then by transposing we obtain $r(1 - e) \leq M(dY)$.

To complete the proof of C_4 for all values of n , we assume it is true for $n = m$ and that G is of class $\mathfrak{G}^{(m+1)}$ on \mathfrak{X}_0 . Then by Lemma 15.3, $d_y G(x, Y; dy)$ and $d_x G(x, Y; dx)$ are of class $\mathfrak{G}^{(m)}$ on $(x_0)_b$ uniformly $(\mathfrak{Y}; \|dy\|)$ and $(\mathfrak{X}; \|dx\|)$, respectively. Next, by Lemma 16.3, the reciprocal of $d_y G(x, Y; dy)$ or $L(x; dy)$ is of class $\mathfrak{G}^{(m)}$ on $(x_0)_b$ uniformly $(\mathfrak{Y}; \|dy\|)$. Then by Lemma 15.2, dY is of class $\mathfrak{G}^{(m)}$ on $(x_0)_b$ uniformly $(\mathfrak{X}; \|dx\|)$, so that, by definition, Y is of class $\mathfrak{G}^{(m+1)}$ on $(x_0)_b$.

18. The unique maximal sheet of solutions containing a given solution. In Theorem 4, the existence of a solution was shown only in a restricted neighborhood of the given solution. However, the region on which the solution is defined can frequently be extended beyond such a neighborhood, by a kind of process of continuation. The maximal sheet of solutions through

a given solution, as defined below, is limited only by boundary points of the region \mathfrak{B}_0 and by points at which the differential $d_v G(x, y; dy)$ has no reciprocal. The theorem is a generalization of one given by Bliss in his *Princeton Colloquium Lectures* (p. 22), and it will be convenient to use a slight modification of his terminology here.

A *sheet of points* in the space $\mathfrak{B} = (\mathfrak{X}, \mathfrak{Y})$ is a set $\mathfrak{B}^{(0)}$ of points $w = (x, y)$ of \mathfrak{B} with the following properties:

(a) for every w_0 of $\mathfrak{B}^{(0)}$ there exist positive constants a and $b \leq a$ such that, no two points of $\mathfrak{B}^{(0)}$ in $(w_0)_a$ have the same "projection" x , and every point x in $(x_0)_b$ is the "projection" of a point w of $\mathfrak{B}^{(0)}$ contained in $(w_0)_a$;

(b) the set $\mathfrak{B}^{(0)}$ is connected. (See § 7.)

A *boundary point of a sheet* $\mathfrak{B}^{(0)}$ is a point not belonging to $\mathfrak{B}^{(0)}$ but every neighborhood of which contains points of $\mathfrak{B}^{(0)}$.

"Sheet" as thus defined corresponds to Bliss's "connected sheet consisting wholly of interior points." This is the only kind we shall have occasion to consider.

Under the hypotheses of Theorem 4, we say that a point $w = (x, y)$ of \mathfrak{B} is an *ordinary point* for the function G if w is in \mathfrak{B}_0 and the differential $d_v G(x, y; dy)$ has a reciprocal. In the contrary case we call w an *exceptional point* for the function G .

We say that a sheet $\mathfrak{B}^{(0)}$ is a *sheet of solutions* of the equation (17.3) in case every $w = (x, y)$ of $\mathfrak{B}^{(0)}$ satisfies (17.3).

THEOREM 5. *If a point $w_0 = (x_0, y_0)$ is an ordinary point for the function G and is a solution of the equation (17.3), then there is a unique sheet $\mathfrak{B}^{(0)}$ of solutions of that equation with the following properties:*

- (C₁) $\mathfrak{B}^{(0)}$ contains w_0 ;
- (C₂) every point of $\mathfrak{B}^{(0)}$ is an ordinary point for the function G ;
- (C₃) the only boundary points of the sheet $\mathfrak{B}^{(0)}$ are exceptional points for the function G .

By Theorem 4, there exists at least one sheet of solutions $\mathfrak{B}^{(1)}$ having properties C_1 and C_2 . Now let $\mathfrak{B}^{(0)}$ be the least common superclass of all such sheets $\mathfrak{B}^{(1)}$. Evidently $\mathfrak{B}^{(0)}$ is a connected set of solutions satisfying C_1 and C_2 . That $\mathfrak{B}^{(0)}$ is a sheet follows from C_2 and Theorem 4. To show that $\mathfrak{B}^{(0)}$ satisfies C_3 , let $(x_1, y_1) = w_1$ be a boundary point of $\mathfrak{B}^{(0)}$ and an ordinary point for G . Since G would then be continuous at (x_1, y_1) , $G(x_1, y_1) = y_*$. Then by Theorem 4 we could extend the sheet $\mathfrak{B}^{(0)}$ to include w_1 , in such a manner that the new sheet satisfies C_1 and C_2 , which contradicts the definition of $\mathfrak{B}^{(0)}$. Now suppose that there is a second

sheet $\mathfrak{B}^{(2)}$ of solutions, having properties C_1, C_2, C_3 . Then $\mathfrak{B}^{(2)}$ is included in $\mathfrak{B}^{(0)}$, and there exists an element w_1 in $\mathfrak{B}^{(0)}$ but not in $\mathfrak{B}^{(2)}$. Since $\mathfrak{B}^{(0)}$ is connected, there exists a continuous function W on \Re to $\mathfrak{B}^{(0)}$ such that $W(r_0) = w_0, W(r_1) = w_1, r_0 < r_1$. By the property C_1 of $\mathfrak{B}^{(2)}$, $W(r_0)$ is in $\mathfrak{B}^{(2)}$. Let r_2 be the least upper bound of the numbers r in the interval $r_0 \leq r \leq r_1$ such that $W(r)$ is in $\mathfrak{B}^{(2)}$. Evidently $W(r_2)$ is a boundary point of $\mathfrak{B}^{(2)}$. But since $W(r_2)$ is in $\mathfrak{B}^{(0)}$, it is an ordinary point for G , which contradicts the property C_3 of $\mathfrak{B}^{(2)}$.

Every sheet of solutions determines a single-valued function $Y(x)$ in a neighborhood of each one of its points. By Theorem 4 each of these functions is of class $\mathfrak{C}^{(n)}$.

UNIVERSITY OF MICHIGAN,
ANN ARBOR, MICH.;
HARVARD UNIVERSITY,
CAMBRIDGE, MASS.