METHODS IN POINT SETS AND THE THEORY OF REAL FUNCTIONS*

BY HENRY BLUMBERG

It is known in psychology that in the succession of psycho-physiological states there operates a so-called law of "facilitation through previous inhibition," which means that an inhibition prepares a reflex or an organic set for more intense activity when the inhibition is removed. Thus tensions produced on hearing discordant notes in music serve to heighten the aesthetic experience following upon the relief of these tensions. Even mathematicians are not quite free from this law; and so it happens, both in the career of individual mathematicians and in the history of the race of mathematicians as a whole, that there may be observed a phase in which the major quest, at least from one point of view, seems to be one of elaborate complexity of pattern, followed by a phase in which the primary concern is for directness and a return to elements. These phases, while perhaps organically antagonistic, as the psycho-physiological law mentioned implies, are, we may suppose, complementary in the larger life.

It is my principal purpose in the present symposium lecture to deal with the latter of these two phases, necessarily characterized by a heightened interest in clearness, intimacy, and economy. The lecture will consist almost entirely of illustrations. These illustrations, chosen for the most part, as is appropriate, from things most familiar to me, will exemplify the emergence of method from procedure guided by this heightened interest in directness and economy. The illustrations will, however, as may be expected, deal not only with this first phase but also with the second, where the interest is in elaboration, one of our very objects being to indicate the passage from the so-called *elementary* to the so-called *involved.*

---

* An address presented at the invitation of the program committee at a Symposium held at the meeting of the Society in Chicago, April 18, 1930. This address presupposes only slight technical knowledge on the part of the reader, and is, for example, intelligible to mathematicians entirely unfamiliar with the theory of point sets.

While we shall not here attempt a general formulation of the principles of method underlying the series of illustrations to be cited, it does appear upon reflection that there are indications of processes of research almost automatic in nature. At any rate, it cannot be doubted that it is of importance for mathematicians to be intimately conversant with the processes that operate in their research and to be in a position consciously to distinguish results that can be obtained almost mechanically after a certain intimacy with the available tools from those that introduce a notable *novum* into our science.

The first illustration, which is familiar to most of us, will exemplify what I call the cue of *restricted choice*; that is to say, sometimes we go certain ways because there isn't much choice. Suppose we want to think about classes or sets without particularizing the character of the elements, as when we think of special sets, for example, of sets of elephants or of points. Can we introduce distinctions in unspecialized sets? There does not seem to be much choice, and so the solution is not hard to find. For since the elements of the sets are in no way to receive special definition, we naturally seek possible distinctions in the abstract *inner* nature of a set, that is to say, in the relationship of a set to itself. Can we introduce a distinction in sets with reference to such an inner relationship? What relationship? The fundamental relationship belonging to the nature of a set is that of containing: the set S contains or does not contain an object e. The distinction will therefore be made with respect to this relationship of containing. Containing what? Since the elements are to be unspecified, we haven't much other choice for the "what" than the set itself. And so we are led to two types of set: (a) sets that contain themselves, like the class of all classes;* and (b) sets that do not contain themselves, like the set of integers, or virtually every set the unsophisticated person is likely to think of. With this distinction, elaboration leads us at once to consider the aggregate of sets of type (a) and the aggregate of sets of type (b). Let us call the first aggregate A and the second, B. The aggregate A, then, consists of sets S such that S contains itself as an element, and B of sets S such that S doesn't contain itself as an element. We ask im-

* We are not entering, in this naïve excursion, upon fine points of rigor.

mediately: Is A in A or in B? There seems to be nothing to prevent A's being in A nor its being in B, a curious and suspicion-arousing indeterminancy. Next, how about B? Suppose it is in A, then it satisfies the relationship S contains S, that is, B contains B; in other words, if B is in A, it is in B. Now suppose B is in B, then it has the characteristic property of the elements S of B, namely that S does not contain S, hence B does not contain B, and therefore B is in A. Thus if B is in B it is in A, and if it is in A it is in B, and we are led to a contradiction with Aristotelian logic. My purpose in repeating this familiar paradox is to indicate the almost mechanical way in which the course of its origin might be conceived.

I shall next describe the simple way in which closed* point sets and other related sets may be thought of as originating. For brevity and simplicity, I shall, generally speaking, confine myself to linear sets, as they are in most cases adequate for illustrating the train of thought to be presented. Now if a linear set C is closed, its complement O with respect to the continuum, that is, the set of points of the continuum not belonging to C, is *open*, by which we mean that every point of O lies in an interval all of whose points lie in O. Now let us say that an interval has the property Q if it lies entirely in O. Then by means of this property Q, we can characterize the sets C and O by saying that every point of O is enclosable in an interval having property Q, and no point of C is so enclosable. In other words, if we define the interval property P as complementary to the property Q, by which we mean that to say that I does not have the property Q is to say that I has the property P, then, in terms of this property P, we can characterize the points of C by saying that every interval containing a point of C has the property P. It follows conversely, and just about as simply, that if P is a given interval property, then the totality of points x such that every interval containing x has the property P is a closed set. Every interval property thus leads to a closed set and every closed set is thus derivable from an interval property. This conclusion shows us how all closed sets are logically derivable from interval properties, but the mere logical connec-

* A set is closed if it contains its limit points. The set of points $1/n$, ($n = 1, 2, \cdots$), is not closed, but it becomes so if the element 0 is added to the set.

tion of ideas or mere logical unifications do not interest us here primarily. It happens, however, that the genesis of closed sets from interval properties is a natural one.

Here are some examples: Let S be a given linear set, and let I have the property P, in notation I^P, if I contains a point of S. In this case, the closed set C attached to the property P is the set of points x such that every interval containing x has the property P, that is, contains points of S. In other words, C equals $S+S'$, where S' is the derived set (=set of limit points) of S. Thus the sum of a set S and its derived set S' is closed. Again, the linear set S being given, let I^P, if I contains an infinite number of points of S; then we conclude that S', the set of limit points of S, is closed; or let I^P, if it contains a non-denumerable subset of S; we then conclude that the set of condensation points* of S is closed. Again, if $f(x)$ is a given function, let I^P, if the saltus of f in $I \geqq$ a fixed number k.† Therefore, the set of points where the saltus of an arbitrary function is greater than or equal to k is a closed set.

The usual presentation in treatises on point sets or real functions seems to invite a separate effort of thought, though here not very great, of course, in showing that a set of a new type is necessarily closed. Here the purpose is shifted to that of listing interval properties, and on a parallel list we shall have closed sets, since for every interval property we are assured of having an associated closed set, though of course, some may be of an especially trivial or uninteresting sort.

In building sets of greater and greater complication, we begin with the interval, which is the simplest set. Using the processes of taking the logical sum and the logical product of sets, and the process of taking the complement of a set with respect to the continuum, we are led next to consider the sum of a sequence of open intervals. Such a sum is apparently an open set. The sum of a sequence of open sets is again an open set, so we do not go beyond open sets by repeated summation of open intervals.

---

* A condensation point of S is a point every neighborhood of which contains a non-denumerable subset of S.

† The saltus of f in I is the upper boundary of f in I minus its lower boundary in I. The saltus of f at a point ξ is the limit of f in a variable interval enclosing ξ and of infinitesimal length.

But we do get a new type of set by taking an infinite product of open sets, a $\prod_1^\infty O_n$, which we write more briefly O_π. The complement of a set of type O_π is of type $\sum_1^\infty C_n$, where C_n designates a closed set, which we similarly shorten to C_σ. That the complement of an O_π is a C_σ follows from simple logic together with the fact that the complement of an open set is closed. These two new types of set, the O_π and the C_σ, may be thought of as coming immediately after the open and the closed sets in the order of increasing structural complexity. Now just as every closed set C may be regarded as attached to an interval property P, so every C_σ may be attached to a sequence of interval properties $\{P_n\}$, $n=1, 2, \cdots$, in the precise sense that $C_\sigma=\sum_1^\infty C_n$, where C_n is the closed set attached to the interval property P_n. We are thus prompted to make a list of sequences of interval properties, especially such sequences as will have attached to them sets C_σ of particular interest. Of course, ingenuity may be requisite in securing such interesting C_σ's, for no general process or theorem can of itself lead automatically to its interesting implications. Still, it does make a difference whether we consciously seek such interesting C_σ's or just stumble upon them, and even excellent mathematicians, in this extremely elementary business of the C_σ's, seem to have secured results in the latter manner, and as a consequence these results were regarded as of some complexity. Also, knowing this simple genesis of the C_σ's, we can without great difficulty add to the interesting C_σ's mentioned in the literature.

I will now cite a number of examples of C_σ's arising from sequences of interval properties.

Let $f(x)$ be a given real function, and let I^{P_n}, if the saltus (defined above) of $f(x)$ in I is greater than $1/n$. If $f(x)$ is discontinuous at ξ, there must be an n such that the saltus of f in every interval enclosing ξ is greater than $1/n$, that is to say, every I containing ξ has the property P_n; ξ thus belongs to the closed set C_n associated with P_n, and therefore to the C_σ attached to the sequence of properties $\{P_n\}$. All the points of discontinuity of $f(x)$ therefore belong to C_σ. On the other hand, if ξ belongs to C_σ, it must belong to some C_n, and therefore, according to the relationship of interval property and its attached closed set, every I containing ξ has the property P_n, that is to say, the saltus of $f>1/n$ in every interval containing ξ, and ξ is therefore a point of discontinuity of f. We conclude that the totality of points of discontinuity of an arbitrary function is identical with C_σ; the set of points of discontinuity of an arbitrary function is of type C_σ. The converse of this result is also true, namely, that if S is any given set of type C_σ, there exists a function $f(x)$ which is discontinuous at every point of S and continuous at every point outside of S.

Let $f(x)$, as before, be an arbitrary real function, and let I^{P_n}, if the d-saltus* of $f(x)$ in I is greater than $1/n$. By reasoning as in the last example of the ordinary saltus, we conclude that the set of points of an arbitrary function where the d-saltus is different from 0, we may call such points, points of d-discontinuity, is of type C_σ. And the converse follows here also. Similar results hold if the negligible set instead of being denumerable, as in the case of the d-saltus, is finite, or of Lebesgue measure 0,† and in all of these cases, as well as in some others, the converse holds also.

Now, while I shall not speak here of the course of reasoning establishing the converse results, I will remark that although to the uninitiated, the proofs of the converse propositions may possibly appear more involved than the train of thought leading to the direct results, yet in fact they call only for a certain skill in workmanship which any student can be rather sure of securing if he applies himself in certain specific ways.

Another example of a theorem that emerges from a suitable definition of a sequence of interval properties is one that involves the notion of *bounded grade*. For instance, for a function $z=f(x, y)$ of 2 variables, we understand by the grade corresponding to 2 points $A=(\xi, \eta)$ and $B=(x, y)$ of the xy-plane the absolute value of the difference of the z's at A and B divided by the

* The d-saltus of $f(x)$ in I is the saltus of f in I on the understanding that denumerable sets are negligible; that is to say, it is the lower boundary of the numbers $s(I-D)$, where $s(I-D)$ is the saltus of f in the set $I-D$, D being an arbitrary denumerable subset of I.

† The Lebesgue exterior measure of a set S, lying in (a, b), is the lower boundary of the length sum of a set of intervals, finite or infinite in number, containing all the points of S as interior points; the Lebesgue interior measure of S equals $b-a$ minus the exterior measure of the complement of S. If the interior measure of S equals the exterior measure of S, this number is the measure of S, and S is then said to be measurable. If the interior and exterior measures are different, S is non-measurable.

distance between A and B; and $f(x, y)$ is said to be of bounded grade at A, if a constant k exists such that for all points B sufficiently near A the grade corresponding to the pair of points A and B is less than k. With this notion of bounded grade, we can, with the aid of a little elementary geometry, define a sequence of properties P_n such that the associated C_σ is precisely the set of points of bounded grade.* And so we have the result that the set of points at which an arbitrary real function is of bounded grade is of type C_σ. The converse theorem holds here also.

I turn next to the consideration of existence proofs in the more elementary parts of the theory of point sets and real functions. There are many methods of proof available for such existence theorems, for example, the method of the Dedekind cut, or of the Bolzano-Weierstrass theorem, etc., which have received attention, for example, in Professor Hildebrandt's lecture on *The Borel theorem and its generalizations.*† I wish here to speak only of two of these methods, that of the descending interval property and the one based on the inductive principle. We say that the interval property P is a descending interval property, if it is such that whenever it holds for an interval I and I is the sum of a finite number of intervals I_ν, $(\nu=1, 2, \cdots, n)$, whether overlapping or not, then it must hold for at least one I_ν. For example, if S is a given infinite set, let us say that I^P, if I contains an infinite subset of S. This property P is then a descending interval property. By means of the fundamental theorem of the Dedekind theory that every Dedekind cut (A, B) in the system of real numbers has either a last in A or a first in B, we can see immediately that if P is a descending interval property, and I has the property P (here we want I to mean a closed interval) then this property P is localized at one point at least of I, *localized* in the precise sense that there exists a point ξ of I such that every neighborhood of ξ contains an interval I having the property P. Now it can be easily shown that all localizable properties can be attached to descending interval properties; that is to say, all theorems

* These properties must now be thought of as referring to the 2-dimensional analog of the linear interval; for example, to the set of interior points of circle.

† This Bulletin, vol. 32 (1926), p. 423.

asserting the existence of a point of a certain character can be deduced from the fact that a descending interval property holding for an interval is necessarily localized at some point of I. Again, the interest here is not that of exhibiting logical deducibility, but rather the genesis of existence theorems from interval properties. This time, however, the property is not unrestricted. as in the case of the genesis of theorems on closed sets, but must be descending, so that we are here invited to make a special list of interval properties, the descending interval properties, knowing that for every property on this list we shall have a corresponding existence theorem. For example, we have seen that if I^P means that I contains an infinite subset of a given set S, then I is descending; therefore, if I^P, P is localized at some point ξ of I, that is to say, is a limit point of S. In other words, the theorem attached to this particular descending property is the Bolzano-Weierstrass theorem, that an infinite set lying in a closed interval has at least one limit point. Again, if $f(x)$ is any given function, let I^P if the variation* of f in I is infinite. Clearly this is a descending interval property. The attached theorem says that if f is of infinite variation in I, there is a point ξ of I such that f is of infinite variation in every interval containing ξ. Or let S be any linear point set, and let I^P if S is non-measurable in I; again, P is a descending interval property, and we conclude that a non-measurable set has at least one point of non-measurability, the meaning of the latter term being given by the very process of attaching an existence theorem to a descending interval property.

It seems that even mathematicians working in point sets or real-variable theory are not always fully conscious of such a natural common genesis of these existence theorems. Indeed, occasionally, as for example in a recent volume of the Fundamenta Mathematicae, we have an article by a well known mathematician containing essentially nothing more than a new example of a descending interval property, and one not remotely accessible. And the proof of the theorem which we would attach to this descending interval property is given in this article in detail.

---

* The variation of f in the interval $I=(a, b)$ is the upper boundary of $\sum_1^n |f(x_\nu)-f(x_{\nu-1})|$ for all possible partitions $a=x_0<x_1<x_2 \cdots <x_n=b$ of (a, b).

We now come to consider the inductive principle. It does not appear to have been remarked in so many words, but it is true that the inductive principle holds just as well for a linear order,* subject to no further restriction, as for the special orders which are called normal, that is to say, orders which, like the set of positive integers, have the property that every subset has a first element.

To state the inductive principle for the general linear order just defined, we need to know the meaning of segment of a linear order. Segment is the abstraction of interval. The subsets S of the linear order O is said to be a segment of O, if S is such that if a and b are two elements of S and c an element of O between a and b, then c belongs to S. S is said to be an initial segment of O if it is such that if b is an element of S, and a an element of O preceding B, then a belongs to S. I want to say specifically that I shall regard the null set 0, that is to say, the set that has no elements at all, as an initial segment of O. We must furthermore define what we mean by an extension of an initial segment S of O. E is said to be an extension of the initial segment $S(\neq 0, O)$, if E is a segment of O having at least one element in common with S and at least one element in common with $C(S)$, the complement of S, that is, the set of elements of O not belonging to S.† If $S=0$, an extension of S means simply any initial segment of S different from 0, and if $S\equiv O$, an extension

* A linear order is an abstraction suggested by the particular order of points on a straight line. A set O conjoined with a relationship of rank—the abstraction of coming before in time or in place or in magnitude—which we denote by the curved inequality sign $\prec$ is said to be a linear order, if the following uniqueness and transitivity properties hold:

Uniqueness: For every pair of elements a, b of O one and only one of the three relationships $a\prec b$, $a=b$, $b\prec a$ holds.

Transitivity: If $a\prec b$ and $b\prec c$, then $a\prec c$.

Examples of linear order are the set of integers arranged in the order of ascending magnitude; the set of real numbers arranged in order of descending. magnitude; the set of positive integers where the rank relationship is defined as follows: the integer $n_1\prec n_2$, if the number of prime factors of n_1 is less than that of n_2; and in case the number of prime factors is the same, then $n_1\prec n_2$, if $n_1<n_2$ in the ordinary sense. The first two orders are not normal, that is to say, not every subset has a first element, not even the set itself, but the last one is.

† This form of definition of extension is chosen with reference to its applicability to the n-fold order; see below.

of S means a final segment of S different from O, final segment being defined similarly to initial segment. The inductive principle for linear order may now be stated as follows:

Let O be a given linear order, and P a given property such that if S is an initial segment of O for all the elements of which P is valid, then there exists an extension E of S such that P is valid for all the elements of E; P must then hold for all the elements of O. In other words, if we call a property of the character just described an *inductive property*, then we can say that an inductive property necessarily holds for all the elements of O.

The agreement that a null set shall count as an initial segment permits us to characterize the inductive property by means of a single mark, instead of two, as is usually done, and this, I think, is not an artificial simplification, but one consonant with the nature of induction. The proof of the inductive principle for ordered sets is, of course, not difficult.

A normally ordered set, as we have said before, is one such that every subset of it has a first element. In particular, if S is an initial segment of a normally ordered set, it is localized by an element e of S, namely the first element of O not in S. In the case of normally ordered sets, it is therefore appropriate to understand by an extension of an initial segment of S merely the set consisting of this first element e not in S. The inductive principle for normally ordered sets therefore takes the form: If the validity of a property P for the elements preceding e, whatever e may be, implies its validity for e, then P holds for every element of O.

If O is the set of points in the linear interval (a, b), then an initial segment S of O is always localized in the sense that there is a last element in S or a first element outside of S. The inductive principle for a linear interval (a, b) then takes the following special form: If the validity of a property P for all the elements $<\xi$ of (a, b) implies its validity for the elements of some interval containing ξ as an interior point, then P holds for all the elements of (a, b). For example, the Borel covering theorem, which also is essentially an existence theorem, is an immediate corollary of this special form of the inductive principle.

One can therefore list existence theorems of the type here considered either by listing descending interval properties or inductive point properties.

The question comes up: Can the inductive principle be applied to n-dimensional space? Or to put it more generally, to an n-fold order? By an n-fold order is understood a system of abstract "points" in which a point e is given by n coordinates, $e=(e_1, e_2, \cdots, e_n)$, but the coordinates are not necessarily magnitudes in the ordinary sense, but elements in n given linear orders, $O_1, O_2, \cdots, O_n$. The way the inductive principle can be extended to an n-fold order is as follows. I shall, for simplicity, speak of the double order. The orders O_1 and O_2, are, then, any two given linear orders, and the double order $O=(O_1, O_2)$ consists of all pairs (e_1, e_2), where e_1 belongs to O_1, and e_2 to O_2. By an initial segment of O we understand a double order $S=(S_1, S_2)$, where S_1 is an initial segment of O_1, and S_2 an initial segment of O_2, and by an extension of S we mean the double order (E_1, E_2), where E_1 is an extension of S_1, and E_2 of S_2. With these definitions of initial segment and extension for the double order, the inductive principle, as stated before for the linear order, takes on definite meaning for the double order, and as so stated, is correct, though the patterning of the proof that seems requisite is somewhat more elaborate than in the case of the linear order.

Mathematical induction, then, is simpler and more powerful than is generally supposed.

An $\aleph_0$-fold order is one in which the elements e have $\aleph_0$ coordinates $e_1, e_2, \cdots$, belonging respectively to $\aleph_0$ linear orders $O_1, O_2, \cdots$. In the time in which I have had opportunity to reflect upon the problem of the inductive principle for the $\aleph_0$-fold order, I have not found for it a satisfactory, valid form.

I now return to the consideration of the descending interval property, and will indicate how a natural elaboration of the idea contained in it leads to results of an apparently considerable degree of sophistication, results which, however, are in reality not far removed in thought process from such simple things as we have so far considered.

Suppose we try to extend the Bolzano-Weierstrass theorem to function space, where by function space we understand here the ensemble of real, continuous functions defined in an interval (a, b). This ensemble by itself does not define a space, since by space we mean more than a mere class of elements, the word *space* implying an interconnection between the elements. One way of introducing such an interconnection in function space is to define the distance between two points of function space, that is to say, the distance between two continuous functions $f_1(x)$ and $f_2(x)$. There are various ways of doing this, but suppose we adopt the following definition: The distance between $f_1(x)$ and $f_2(x)$ is max $|f_1(x)-f_2(x)|$ as x ranges over (a, b). In terms of distance, we can define the *sphere* in function space having $c(x)$ as center and r as radius as the set of points $f(x)$ whose distance from $c(x)$ is $<r$. Sphere thus signifies *interior of a sphere.* In terms of distance we can also define limit point of a set: $f(x)$ is said to be a limit point of S if every sphere containing $f(x)$ contains an infinite number of points of S; sphere thus takes the place of interval in the linear case. Now to extend the Bolzano-Weierstrass theorem to function space is to ask under what general conditions an infinite set S has a limit point. In the linear case, the only condition attached to S is that S be bounded, and clearly the Bolzano-Weierstrass theorem becomes false if this condition is dropped. This condition of boundedness is surely necessary also in the case of function space, that is to say, S must be, as we say, uniformly bounded, which means that there is a fixed constant k such that $|f(x)|<k$ for all f's of S and all x's of (a, b). Is this condition of uniform boundedness also sufficient? Clearly not, as the example $(a, b)=(0, 1)$, $S=\{x^{1/n}\}$, $n=1, 2, \cdots$, shows.

Let us examine the simplest case of a subset S of function space which does possess a limit point, namely the case of a sequence of functions $f_n(x)$ having only a single limit point $f(x)$. Every sphere containing $f(x)$ contains all but a finite number of the $f_n(x)$. Let R be a sphere with center $f(x)$ and radius ϵ. Then if (α, β) is a subinterval of (a, b), the saltus in (α, β) of every $f_n(x)$ which lies in R cannot exceed the saltus of $f(x)$ in (α, β) by more than 2ϵ; and since outside of R we have only a finite number of elements of S, we conclude from the fact that a continuous function defined in a closed interval is uniformly continuous, that the set S is *equicontinuous.* This means that for every $\epsilon>0$, there exists a $\delta>0$ such that $|f_n(x_1)-f_n(x_2)|<\epsilon$ for all n's and for all pairs of numbers x_1, x_2 such that $|x_1-x_2|<\delta$. Now upon slight reflection it can be seen that if S is any given infinite subset of function space, instead of just a sequence of functions with a unique limit, then, though S is restricted to being uniformly bounded, there will always be infinite subsets of it without limit points unless this condition of equicontinuity is imposed, equicontinuity really signifying nothing more than another type of boundedness. It can be shown that the method of the descending interval property, for example, can be readily extended to function space as we have defined it; and by means of this method, we can prove that if S is an infinite set of points of function space subject to the two boundedness conditions, namely uniform boundedness and equicontinuity, then S necessarily has a limit point. And the argument is similar to the one for the Bolzano-Weierstrass theorem. Following this direction of elaboration, we can go on to measurable functions,* for example.† The distance between two measurable functions $f_1(x)$ and $f_2(x)$ may then be defined according to the following rough idea: The two functions are not far away from one another if they do not differ much from each other except possibly in a set of slight measure. This rough idea when rigorously stated takes the following form: Let ρ be a number such that $|f_1(x) - f_2(x)| < \rho$ except in a set of measure $< \rho$; ρ may then be regarded as an upper estimate for the distance between f_1 and f_2. Then if d is the lower boundary of all such ρ's, d is the distance between f_1 and f_2. With this definition of distance between two measurable functions, the convergence of the sequence of measurable functions $f_n(x)$ to the function $f(x)$ may be seen to mean what is usually termed *convergence on the average*. Now the analogy with continuous functions readily shows us how to extend, in a natural manner, the notions of uniform boundedness and equicontinuity to this space of measurable functions. The method of the descending interval property is valid, and by means of it the Bolzano-Weierstrass theorem can be proved.

Lastly, in the matter of existence theorems, I may mention that the line of elaboration here sketched carries through, in a

---

* $f(x)$ is said to be measurable, if for every real number k, the set $E_{f>k}$, which is the set of x's such that $f(x)>k$, is measurable. It follows that the sets $E_{f\geqq k}$, $E_{f<k}$, $E_{k<f<l}$, $E_{k\leqq f\leqq l}$, where l is also an arbitrary real number, are measurable.

† See Fréchet, *Sur les ensembles compacts de fonctions mésurables*, Fundamenta Mathematicae, vol. 9(1927), p. 25.

certain sense, to a space consisting of functions which are entirely unconditioned, this by means of a general method for such extensions which I shall speak of at the close of the lecture.

I shall now speak of a theorem on unrestricted functions of two variables which yields a theorem on arbitrary functions of one variable every time we are given an interval function, that is to say, a relation which attaches to every interval a number, like the saltus of a given function in an interval I. Let, then, $z=f(x, y)$ be a given real function of two variables, and s any straight line in the xy-plane,—I am taking the simplest case. Suppose we consider the behavior of $f(x, y)$ as (x, y) approaches s from the same side of s along two given directions $d_1 \neq d_2$. If P is a point of s, we denote by u_{Pd} the lim sup $f(x, y)$ as (x, y) approaches P along d; similarly l_{Pd} will denote the lim inf $f(x, y)$ as (x, y) approaches P along d. Now suppose P is a point of s such that $l_{Pd_1} \geqq k$, where k is a given number; then there is a segment Q_nP having the direction d_1 such that $f(x, y) > k-1/n$ at all points of Q_nP. From Q_n draw Q_nR_n in the direction d_2 meeting s in R_n. Let us say that a point V of s has the character (k, n) if we can find a point U at distance $<1/n$ from it such that UV has the direction d_2 and $f(x, y) > k-1/n$ at U. Clearly if Q_nP is taken small enough—and to have the functional value at all the points of it $>k-1/n$ we can take it as small as we please—every point in the closed interval PR_n except possibly P will have the character (k, n). Now a point V of s which has the character (k, n) for all positive integers n is a point such that as near it as we please we can find a U, with UV in the direction d_2, such that $f > k-1/n$ at U. This means that $u_{Vd_2} \geqq k$. There is a theorem of Young—proved almost immediately from the fact that the linear continuum contains a set, namely that of the rational numbers, which is denumerable and dense in it—which goes as follows: Let $J = \{I\}$ be a given set of closed intervals lying on a straight line; and e such that it is an end point of at least one I of J, but an interior point of no I of J. Then the totality of such points e is at most denumerable. We have remarked before that if Q_nP is small enough, every point of PR_n with the possible exception of the end point P is of character (k, n). By looking into the situation a bit more, we can, with the aid of this theorem of Young, arrive at the following conclusion: If k is a given number, then the points P of s such that

$l_{Pd_1} \geqq k$ and $u_{Pd_2} < k$ constitute a set which is at most denumerable. Another easy step, depending again upon the denumerability and denseness of the rational numbers, and we come to the following theorem:* Let $z = f(x, y)$ be a given arbitrary function of two variables; s any straight line in the (x, y) plane; and d_1 and d_2 two directions of approach to s from the same side of it. Then for every point P of s, with the possible exception of a denumerable number, we have $u_{Pd_1} \geqq l_{Pd_2}$. In other words, if we call the pair of numbers (l_{Pd}, u_{Pd}) the interval of approach of $f(x, y)$ at P along the direction d, then our result can be stated in this way: At every point of s, with the possible exception of a denumerable set, the interval of approach along d_1 overlaps or abuts the interval of approach along d_2. I shall call this result the *theorem on approach*. Clearly this theorem becomes false if d_1 and d_2 are taken on different sides of s, for then d_1 does not cross d_2, and it is precisely on the crossing of two different directions on the same side of a line that the validity of the theorem on approach depends.

The theorem on approach is exhaustive in a certain sense, but I shall not stop here to explain this further.

Now suppose, in particular, the straight line s is the 45° line of the xy-plane; and $f(x, y)$ is a symmetric function in its arguments. Then as far as the functional values of $f(x, y)$ are concerned, it is the same to approach s vertically downward as to approach it horizontally to the left. The upshot of this fact is the following corollary of the theorem on approach: If $f(x, y)$ is a given symmetric function, then for all x, with the possible exception of a denumerable set, we have $\limsup f(x, x \pm 0) \geqq \liminf f(x, x \mp 0)$.

Now to have a symmetric function $f(x, y)$ is to have a function of an interval $I = (x, y)$, if we only understand, as is natural, that the interval (y, x) is not different from the interval (x, y), so that this corollary of the theorem on approach yields a theorem for every interval function. Again we are invited to make a list,—this time of interval functions—and in a natural way, we rediscover a variety of results which have appeared from time to time in the literature without revealing a common

* See Fundamenta Mathematicae, vol. 16 (1930), p. 17.

origin, and we also obtain new applications of interest. I shall give some examples.

If the symmetric function of our corollary is the difference quotient $(f(x)-f(y))/(x-y)$, we obtain the theorem of G. C. Young, that the upper right (left) derivative of an arbitrary function is greater than or equal to the lower left (right) derivative except possibly at a denumerable number of points.

Another illustration. Let $f(x)$ be a given real function. If (ξ, η) is a real interval, and p and q two real numbers such that $p>q$, let $u(f;\xi,\eta;p,q)$ represent the upper boundary in (ξ, η) of those values of $f(x)$ that lie between p and q. By using this function of the interval (ξ, η) for all possible pairs of rational numbers $p>q$, we arrive at the following theorem of W. H. Young.

If $f(x)$ is an arbitrary real function, then there exists a set D, which is at most denumerable, such that if ξ is a point not in D, every point in the (x, y) plane on the line $x=\xi$ which is a limit of the points of the *curve* $y=f(x)$ from the right (left) is also a limit of these curve points from the left (right).

We can readily obtain, from our point of view, interesting extensions of this theorem of Young.

One other example. Let (ξ, η) and (p, q) have the same meaning as in the preceding example, and let r be a real number between 0 and 1. By $u(f;\xi,\eta;p,q;r)$ we understand the upper boundary of the same set of points as before, except that now, in computing this upper boundary, we are privileged to neglect any subset in the interval (ξ, η) which is of relative exterior measure $<r$; that is to say, we may neglect a set of Lebesgue exterior measure $<r(\eta-\xi)$; $u(f;\xi,\eta;p,q;r)$ is thus, to be more accurate, the lower boundary of the set of upper boundaries corresponding to all possible such negligible subsets of exterior measure $<r(\eta-\xi)$. By considering this function of the interval (ξ, η) for all possible pairs of rational numbers $p<q$, and for r varying and approaching 0, we secure a series of theorems for unrestricted functions, one of which is a theorem due to Kempisty. For its statement we need to know the meaning of upper metric limit and lower metric limit of a function $f(x)$ at the right or left of a point ξ. By the upper metric limit of $f(x)$ at the right of ξ, we mean the lower boundary of all numbers a

such that the (exterior) metric density* of the set of x's for which $x > \xi$ and $f(x) > a$ [we denote this set by $E(x > \xi; f(x) > a)$] is 0 at ξ; similarly, the lower metric limit at the right of ξ is the upper boundary of all numbers a such that the (exterior) metric density of the set $E(x > \xi; f(x) < a)$ is 0 at ξ; and in the same way, we define the upper metric limit and lower metric limit at the left of ξ. The theorem of Kempisty may now be stated as follows: For every real function, the upper metric limit at the right (left) is greater than or equal to the lower metric limit at the left (right) except possibly at a denumerable number of points.

I shall next speak of certain properties of real functions which can be obtained, according to a definite procedure, from certain properties of linear sets. I shall take first a simple example for which the reasoning can be presented, I think, even in such a lecture as this without burdening the listeners. Suppose S is any given linear set. Then we shall say that a point x of S is densely approached by S, if there is an interval I containing x in which S is dense.† It is easy to see that those points of S which are not densely approached by S constitute a nowhere dense set.‡ For suppose N is this subset of points of S at which we do not have dense approach. A fortiori, then, no point of N is densely approached by N; that is to say, there is no interval I containing a point of N in which N is dense, in other words, every interval containing a point of N contains a subinterval in which there are no points of N, which amounts to saying that N is a nowhere dense set. Now obviously, the sum of a sequence of nowhere dense sets need not be nowhere dense. But in the sequel we shall be interested only in such types τ of set that if $S_1, S_2, \cdots, S_n, \cdots$ is a sequence of sets of type τ, then the sum of the sequence must also be of type τ.§ The type of nowhere dense set therefore does not suit our present purpose.

* The (exterior) metric density of a set S at a point x is the limit (if it exists) of $m_e(SI)/l$, where $m_e(SI)$ means the exterior Lebesgue measure of the set of points common to S and I, as the interval I, containing x, varies in any manner with its length l infinitesimal.

† S is dense in I if every subinterval of I contains points of S.

‡ S is nowhere dense in the continuum if there is no interval in which it is dense; that is, every interval contains a subinterval having no points of S.

§ A more general formulation can be made in which this property is not required.

However, we can help ourselves if we define an exhaustible set as a set which is representable as the sum of a sequence of nowhere dense sets. It then follows from the fact that a double sequence can be rearranged as a single sequence, that the sum of a sequence of exhaustible sets is exhaustible. I will remark in passing that an exhaustible set acts like a set poor in elements, in the descriptive, that is to say, non-metric portion of the theory of point sets. An exhaustible set never exhausts the continuum, that is to say, the complement of an exhaustible set is not empty. In fact the complement of an exhaustible set is itself not exhaustible; and it is easy to show this, but I shall not stop to do so.

Returning now to the given linear set S, we have seen that the points of S which are not densely approached by S constitute a set which is nowhere dense. We can now say that these points constitute a set which is exhaustible, and this is exactly what we want to say, because exhaustibility reproduces itself upon infinite summation.

This property of every linear set S, namely the exhaustibility of the set of points of S at which the approach is not dense, we shall now utilize to derive a property of an unrestricted function. We shall say that the point $(\xi, f(\xi))$ is densely approached by the curve $y=f(x)$, or that the approach of $f(x)$ is dense at ξ, if for every pair of real numbers k, l such that $k<f(\xi)<l$, the set $E_{k<f<l}$ has dense approach at ξ. Evidently, in the definition of dense approach of a function at a point, we may confine ourselves entirely to rational numbers k, l. We then have, in all, only a denumerable number of sets $E_{k<f<l}$ to deal with, and since each of these linear sets has dense approach at all of its points except possibly at the points of an exhaustible set, and since there are in all only a denumerable number of such linear sets, we conclude that the totality of exceptional points is the sum of a sequence of exhaustible sets and therefore itself exhaustible. This yields the following theorem.

An arbitrary function has dense approach everywhere except possibly at the points of an exhaustible set.

The converse result, as may be briefly shown, holds also; that is to say, if E is an arbitrarily given, exhaustible set, a function exists which does not have dense approach at the points of E but has dense approach at every point of the complement of E.

Now underlying the method we have here sketched for obtaining the theorem on general functions concerning dense approach, there is a procedure for passing from properties of sets to properties of functions. We may formulate this procedure, in one aspect, as follows: let R be a relationship of a set S to a point x of it of such a sort that if S is any given set, then S has the relationship R to every point of it with the possible exception of a set E of type τ. Moreover, we assume that the sum of a sequence of sets of type τ, is of type τ, and additionally, that a subset of a set of type τ is of type τ. Then with this relationship R we associate as follows a relationship R' of a function $f(x)$ to a point ξ. We say $f(x)$ has the relationship R' to ξ if for all $k<f(\xi)<l$ the set $E_{k<f<l}$ has the relationship R to ξ. With this understanding, we can say that $f(x)$ has the relationship R' to x at all points x with the exception of a set of points of type τ.

From this point of view we derive such results as the following:

An arbitrary function is inexhaustibly approached, in a sense that can be readily surmised, everywhere except possibly at the points of an exhaustible set.

An arbitrary function is quasi-continuous almost everywhere,* that is to say, everywhere except possibly in a set of Lebesgue measure zero.

This result can be directly extended to many-valued functions. We may, indeed, allow $f(x)$ to take on k_x values at x, this number k_x being any number between 1 and c, the cardinal number of the continuum, and varying with x in any manner. If $f(x)$ is such an arbitrary, many-valued function, then there exists a set Z of Lebesgue measure zero such that if ξ is a point not belonging to Z, and $(\xi, f(\xi))$ is any one of the points of the curve $y=f(x)$ on the line $x=\xi$, then $y=f(x)$ is quasi-continuous at ξ.

Going on a step from this point, we can reach the following result, but we must first explain what we mean by positive approach and full approach of a curve $y=f(x)$ at a point (ξ, η), which, this time, may or may not belong to the curve itself. The curve $y=f(x)$ is said to approach the point (ξ, η) positively

---

* The function f is quasi-continuous at ξ, if the set $E_{k<f<l}$ is, for every pair of real numbers $k<f(\xi)<l$, of exterior metric density 1 at ξ.

if for every $k<\eta<l$, the set $E_{k<f<l}$ is not of metric density 0 at ξ; and the approach is full at (ξ, η) if for every $k<\eta<l$ the metric density of $E_{k<f<l}$ equals 1. Then we can state the following theorem. For every function $f(x)$, the points ξ of the sort that there is even a single η such that (ξ, η) is positively but not fully approached by the curve $y=f(x)$ constitute a set of measure zero.

I shall speak lastly of a theorem that enables us in various interesting cases to derive properties of unrestricted functions from properties of measurable functions; and since measurable functions are in certain essential respects not far removed from continuous functions, we thus have a challenge, every time a property of continuous functions is presented, to extend it to arbitrary functions.

The theorem I have in mind may be formulated as follows. *If $f(x)$ is any given function, there exist two measurable functions $m_1(x)$ and $m_2(x)$ such that the set of points x for which $f(x)<m_1(x)$ is of measure zero; likewise the set of points x for which $f(x)>m_2(x)$ is of measure zero. Moreover, at every point of both $y=m_1(x)$ and $y=m_2(x)$, with the exception of a set of measure zero, the metric approach of $y=f(x)$ is full.*

We shall call the functions $m_1(x)$ and $m_2(x)$ the measurable boundaries of $f(x)$; and the theorem just stated we shall call the theorem on the measurable boundaries.

More can be said, in a similar sense, concerning the nature of $f(x)$ between its measurable boundaries, but I shall not stop for this here.

The theorem on the measurable boundaries gives us much insight into the structure of an arbitrary function, and it is for this reason that it is of considerable utility. For instance, there is a theorem of Vitali which says that every measurable function is equal to a function of class two except at points of a set of measure 0.* This theorem has been generalized to arbitrary functions by Sierpiński and Saks† by an argument of some length. By means of the theorem on the measurable boundaries,

* A continuous function is of class zero; if f is not continuous but the limit of a sequence of continuous functions, it is of class one; if not of class one, but the limit of a sequence of functions of class one, it is of class two; and so on.

† Fundamenta Mathematicae, vol. 11 (1928), p. 105.

we can derive this generalization of Vitali's theorem almost at a glance. Also the theorem of Bolzano-Weierstrass generalizes in this way, under a certain form, to unconditioned functions. Another example is that of the rather exhaustive results of Denjoy on derivatives, first proved for continuous functions, and later extended by G. C. Young to measurable functions. When these results of Denjoy are extended to unrestricted functions, we secure a result of apparently great profundity but really not as enormously removed, as one might from the outside suppose, from simple things. This result may be stated as follows: At every point $P=(\xi, f(\xi))$ of an arbitrary curve $y=f(x)$, with the exception of a set of ξ's of measure zero, the directions of approach to P along the curve lie either in a $0°$ angle—case of the existence of the derivative; or in an angle of $180°$, but not less; or in an angle of $360°$, but not less.*

I will remark in closing that there are ways of search in mathematics very different from those indicated in the present lecture. For example, an eminent mathematician once told me that he had obtained nearly all of his important results "at the point of the pen," as he put it; that is to say, not by perseveringly seeking a conceptual envisagement of utmost simplicity, but, as it were, allowing his mathematico-psychophysiological set to determine the course of action of his brain and hand, just as musical compositions have been sometimes created. Most of the time, he said, nonsense resulted, but now and then he was guided to beautiful things.

There are no a priori reasons for assuming that there are not still other quite different modes of search; there are, indeed, positive indications of such.

This lecture was concerned with the mode in which discursive thinking by means of concepts attains its end in mathematics. A similar analysis, not without promise, is conceivable for the mode that operates "at the point of the pen," and for the study of this mode, access to private notes would be invaluable.

Surely illustrations of method similar to those here selected can be found in other mathematical fields. To become aware of

---

* See S. Saks, *Sur les nombres dérivés des fonctions*, Fundamenta Mathematicae, vol. 5 (1924), p. 98, where the extension of Denjoy's results to arbitrary functions is made by means of an independent course of reasoning.

these methods is, of course, not in itself an end, but a means toward increasing one's agility in the use of processes, and perhaps even a guide toward new possible modes of search. To achieve such consciousness of process should be of particular interest to those just entering upon the serious pursuit of our fascinating science. In this connection, it may be helpful to reflect that the embarrassment which mathematics produces in certain quarters comes rather from the extensiveness of its domain than from the inherent difficulty of its processes.

The present study points to the possibility of higher unifications in mathematics from the point of view of process rather than content or logical concept; in other words, unifications not from the point of view of conceptual structure but from the point of view of a behavioristic psychology.

Ohio State University