The Generalized Weierstrass Approximation Theorem
Author(s): Marshall H. Stone
Source: *Mathematics Magazine*, Vol. 21, No. 5 (May - Jun., 1948), pp. 237-254
Published by: Taylor & Francis, Ltd. on behalf of the Mathematical Association of America
Stable URL: https://www.jstor.org/stable/3029337
Accessed: 19-12-2024 03:12 UTC

THE GENERALIZED WEIERSTRASS APPROXIMATION THEOREM

by Marshall H. Stone

(Continued from March-April issue)

5. *Extension to Complex Functions.* It is natural to consider the extension of the preceding results to the case of complex-valued functions. The fact that the complex numbers are not ordered is an obstacle to the introduction of lattice operations for complex-valued functions. Accordingly the results of §2 do not lend themselves to extension in the desired sense, unless they are first expressed in terms of the operation of forming the absolute value. However it is easy to see that the complex-linear operations (addition and multiplication by complex numbers) and the operation of forming the absolute value do not work well enough together for us to obtain any very interesting or useful extension to the complex case. Matters appear quite differently when we consider the linear ring operations. In fact, we find here that extremely interesting new possibilities, quite beyond the scope of the present inquiry, are immediately opened up. For example, the theory of analytic functions can be considered as an answer to the question, "If X is a bounded closed subset of the complex plane, and $\mathfrak{X}_0$ is the family of all polynomials in the complex variable z, what functions can be uniformly approximated on X by functions in $\mathfrak{X}_0$?" As this observation clearly suggests, a full investigation of the complex case would be both difficult and rewarding. In order to limit ourselves to considerations of the kind already met in the preceding sections, we shall include the operation of forming conjugates along with the linear ring operations in our examination of the complex case. Accordingly, if $\mathfrak{X}$ is the family of all bounded continuous complex-valued functions on a topological space X and $\mathfrak{X}_0$ is an arbitrary subfamily of $\mathfrak{X}$, we designate by $\mathfrak{U}(\mathfrak{X}_0)$ the family of all those functions which can be obtained from $\mathfrak{X}_0$ by the linear ring operations, the operation of forming conjugates, and uniform passages to the limit. It is easily seen that $\mathfrak{U}(\mathfrak{X}_0)$ is obtainable by using first the algebraic operations and then a single passage to the limit; and that $\mathfrak{U}(\mathfrak{X}_0)$ is closed under all the operations permitted. If we designate the conjugate by means of a dash, so that $\bar{f}$ is the function whose value $\bar{f}(x)$ at x is equal to the conjugate of the complex number $f(x)$, we can define two related operations, namely those of forming the real part and the imaginary part of f, by the equations

$$\Re f = \tfrac{1}{2}(f + \bar{f}) \qquad \Im f = 1/2i(f - \bar{f})$$

from which the relations

$$f = \Re f + i\,\Im f, \qquad \bar{f} = \Re f - i\,\Im f$$

follow directly. The functions $\Re f$ and $\Im f$ are real-valued continuous functions which belong to $\mathfrak{U}(\mathfrak{X}_0)$ whenever f is in $\mathfrak{X}_0$. It is easy to see that $\mathfrak{U}(\mathfrak{X}_0)$ can be obtained in the following manner: we first form the family $\mathfrak{Y}_0$ of all real-valued functions expressible as $\Re f$ or $\Im f$ where f is in $\mathfrak{X}_0$; we then form the family $\mathfrak{B}(\mathfrak{Y}_0)$ of all those (real) functions which can be obtained from $\mathfrak{Y}_0$ by the real linear ring operations and uniform passage to the limit; and finally we find $\mathfrak{U}(\mathfrak{X}_0)$ as the family of all functions $f + ig$ where f and g are in $\mathfrak{B}(\mathfrak{Y}_0)$. In view of this observation we can carry Theorem 5 and its corollaries over to the complex case without any further difficulty. The results will be given without further discussion, as follows.

Theorem 10: *Let X be a compact space, $\mathfrak{X}$ the family of all continuous complex functions on X, $\mathfrak{X}_0$ an arbitrary subfamily of $\mathfrak{X}$, and $\mathfrak{U}(\mathfrak{X}_0)$ the family of all functions (necessarily continuous) generated from $\mathfrak{X}_0$ by the linear ring operations, the operation of forming the conjugate, and uniform passage to the limit. Then a necessary and sufficient condition for a function f in $\mathfrak{X}$ to be in $\mathfrak{U}(\mathfrak{X}_0)$ is that f satisfy every linear relation of the form $g(x) = 0$ or $g(x) = g(y)$ which is satisfied by all functions in $\mathfrak{X}_0$. If $\mathfrak{X}_0$ is a closed linear subring of $\mathfrak{X}$ which contains $\bar{f}$ together with f — that is, if $\mathfrak{X}_0 = \mathfrak{U}(\mathfrak{X}_0)$ — then $\mathfrak{X}_0$ is characterized by the system of all the linear relations of this kind which are satisfied by every function belonging to it. In other words, $\mathfrak{X}_0$ is characterized by the partition of X into mutually disjoint closed subsets on each of which every function in $\mathfrak{X}_0$ is constant and by the specification of that one, if any, of these subsets on which every function in $\mathfrak{X}_0$ vanishes.*

Corollary 1: *In order that $\mathfrak{U}(\mathfrak{X}_0)$ contain a non-vanishing constant function it is necessary and sufficient that for every x in X there exist some f in $\mathfrak{X}_0$ such that $f(x) \neq 0$.*

Corollary 2: *If $\mathfrak{X}_0$ is a separating family for X, then $\mathfrak{U}(\mathfrak{X}_0)$ either coincides with $\mathfrak{X}$ or is, for a uniquely determined point x_0, the family of all functions f in $\mathfrak{X}$ such that $f(x_0) = 0$. If, conversely, $\mathfrak{X}$ is a separating family for X and $\mathfrak{U}(\mathfrak{X}_0)$ either coincides with $\mathfrak{X}$ or is the family of all those f in $\mathfrak{X}$ which vanish at some fixed point x_0 in X, then $\mathfrak{X}_0$ is a separating family.*

In the case of complex-valued functions, the definition of an ideal remains the same: a non-void subclass $\mathfrak{X}_0$ of $\mathfrak{X}$ is said to be an ideal if $\mathfrak{X}_0$ contains $f + g$ whenever it contains both f and g, and $\mathfrak{X}_0$ contains fg whenever it contains f, g being an arbitrary function in $\mathfrak{X}$. Clearly, an ideal is closed under the linear ring operations, since multiplication by a complex number is equivalent to multiplication by a constant complex-valued function. Now if an ideal $\mathfrak{X}_0$ is closed under uniform passages to the limit we can show that $\mathfrak{X}_0$ contains $\bar{f}$ together with f. For an arbitrary function f in $\mathfrak{X}_0$ we define a function g_n by putting $g_n(x) = \bar{f}(x)\,|f(x)|^{1/n}/f(x)$ when $f(x) \neq 0$ and $g_n(x) = 0$ when $f(x) = 0$. If x is any point in X and x_0 any point such that $f(x_0) = 0$ we have

$$|g_n(x)| = |f(x)|^{1/n}, \qquad |g_n(x) - g_n(x_0)| = |f(x) - f(x_0)|^{1/n},$$

whence g_n is bounded on X and continuous at x_0. On the other hand, it is evident that g_n is continuous at any point x where $f(x) \neq 0$. Thus g_n is a function in $\mathfrak{X}$ and fg_n a function in the ideal $\mathfrak{X}_0$. The function $h_n = |\bar{f} - fg_n| = |\bar{f}(1 - |f|^{1/n})| = ||f| - |f|^{1+1/n}|$ is evidently a real continuous function on X which converges as $n \to \infty$ to the function everywhere equal to zero. for any x it is easily verified that $h_n(x)$ is a non-increasing sequence. Accordingly Corollary 1 to Theorem 1 shows that the convergence is uniform. It follows that fg_n converges uniformly to $\bar{f}$ and hence that $\bar{f}$ is in the closed ideal $\mathfrak{X}_0$. As a result we can state without further preliminaries the following extension of Theorem 8 from the real to the complex case, essentially due to Šilov.

Theorem 11: (Šilov [3]). *Let $\mathfrak{X}$ be the linear ring of all the continuous* *complex functions on a compact space* X, *let* $\mathfrak{X}_0$ *be an arbitrary non-void subfamily of* $\mathfrak{X}$, *let* X_0 *be the closed set of all those points* x *at which every function* f *in* $\mathfrak{X}_0$ *vanishes, and let* $\mathfrak{Y}_0$ *be the family of all those functions* f *in* $\mathfrak{X}$ *which vanish at every point of* X_0. *Then* $\mathfrak{Y}_0$ *is the smallest closed ideal containing* $\mathfrak{X}_0$; *and* $\mathfrak{X}_0$ *is a closed ideal if and only if* $\mathfrak{X}_0 = \mathfrak{Y}_0$. *A closed ideal* $\mathfrak{X}_0$ *is characterized by the associated closed set* X_0; *in particular,* $\mathfrak{X}_0 = \mathfrak{Y}_0 = \mathfrak{X}$ *if and only if* X_0 *is void.*

Proof: It is obvious that $\mathfrak{Y}_0$ is an ideal. If a closed ideal $\mathfrak{X}_1$ contains $\mathfrak{X}_0$ then by virtue of what we have just proved above it is a linear subring which contains $\bar{f}$ along with f; and hence a simple argument based on Theorem 10 shows that $\mathfrak{X}_1$ contains $\mathfrak{Y}_0$. The remainder of the theorem follows then in a familiar way.

6. *The Extension to Locally Compact Spaces.* A natural question arises as to the possibility of relaxing the topological conditions imposed hitherto upon the space X. A thorough examination of this question would take us too far afield. Suffice it to say that a great deal of light can be thrown on this question by applying the theory of compactification developed by the writer in [1] and by Čech in a later paper [4]. In fact, it can be said that this theory allows us to solve the problem of approximation in the family $\mathfrak{X}$ of all bounded continuous real (or complex) functions on an *arbitrary* topological space X — in exactly the same sense that the problem has been solved above for compact X. The essence of the method alluded to in these remarks is to replace X by a suitable compactification X^*, extending every function in $\mathfrak{X}$ over the compact space X^* without sacrifice of its continuity. There is however, one very special instance of sufficient immediate interest for us to pay it some attention here. This is the case where X is a locally compact space.

By definition a space is locally compact if every point of the space is interior to some compact subset of the space. Typical examples of spaces which are locally compact without being compact are afforded by the euclidean spaces of n dimensions. If X is a locally compact space which is not compact it can be compactified, as is well known, by the adjunction of a single point. Specifically, we adjoin an element x_∞ to X, obtaining the set X^*, and we define a subset U^* of X^* to be open if it is a subset of X and is open in X, or if it is the complement of a *closed* compact subset of X (note that, while a closed subset of a compact space is compact, a compact subset of a topological space is not necessarily closed unless it is a Hausdorff space!). The totality of open sets in X^* is easily verified to have the properties normally required: X^* and its void subset are open; the union of any family of open sets is open; and the intersection of any finite family of open sets is open. Moreover, X^* can be shown to be compact, as follows: in any family of open sets whose union is X^* we can find one open set containing x_∞; its complement X_0 is compact and is contained in the union of the remaining open sets in the family; but then a finite number of the latter must, because X_0 is compact, have a union containing X_0.

In order that a real function f defined on X should agree there with a function f^* defined and continuous on X^*, it is necessary and sufficient that the function f, in addition to being continuous on X, should satisfy the inequality $|f(x) - f(y)| < \epsilon$ for all x and y outside a suitable closed compact set X_ϵ. The necessity of the condition is obvious: when f^* exists, there is an open set U_ϵ^* containing x_∞ such that for all x and y in it $|f^*(x) - f^*(x_\infty)| < \epsilon/2$, $|f^*(y) - f^*(x_\infty)| < \epsilon/2$, and hence $|f(x) - f(y)| = |f^*(x) - f^*(y)| < \epsilon$; the appropriate set X_ϵ is therefore the complement of U_ϵ^*. The sufficiency of the condition is easy to prove aside from the determination of the value which should be assigned to f^* at x_∞. Let Z be the closure of the set of all real numbers $\zeta = f(x)$, where x is in X and is restricted to lie in a fixed open subset of X^* containing x_∞. Since X is not compact the intersection of any finite number of such open sets must contain points of X, and the corresponding sets Z therefore have a common point. Thus any finite number of the sets Z will have a common point. There exists a set Z of diameter not exceeding an arbitrarily prescribed positive number ϵ, since we may determine a set X_ϵ in accordance with the assumed condition and may take Z as the set corresponding to the open set complementary to X_ϵ: the relation $|f(x) - f(y)| < \epsilon$ for all x and y outside X_ϵ implies that the diameter of Z does not exceed ϵ. Any such set Z, being closed and bounded, is compact. Hence there exists a unique real number ζ_∞ common to all the sets Z. We now put $f^*(x) = f(x)$ for x in X and $f^*(x_\infty) = \zeta_\infty$. Obviously the function so defined is continuous at every point of X. To show that it is continuous at x_∞, we prescribe $\epsilon > 0$ arbitrarily, determine a corresponding set $X_{\epsilon/2}$ by virtue of the assumed condition, and take Z as the set associated with the open set $U_{\epsilon/2}^*$ complementary to $X_{\epsilon/2}$. Obviously there is a point y in $U_{\epsilon/2}^*$ and in X such that $|f(y) - \zeta_\infty| < \epsilon/2$; and at the same time $|f(x) - f(y)| < \epsilon/2$ for all x and y in $U_{\epsilon/2}^*$. Hence $|f^*(x) - f^*(x_\infty)| = |f(x) - \zeta_\infty| < \epsilon$ for all x in $U_{\epsilon/2}^*$ and in X. Consequently f^* is continuous at x_∞, as we desired to show.

Since the extension of a function f satisfying the given condition is uniquely determined, it is clear that the study of the continuous functions on X which do satisfy the condition is equivalent to the study of the continuous functions on X^*. Hence if we denote the totality of such functions as $\mathfrak{X}$ and the totality of their extensions by $\mathfrak{X}^*$, the approximation theorems in $\mathfrak{X}$ are translatable into approximation theorems in $\mathfrak{X}^*$ and vice versa. In making the indicated translation it is frequently convenient to have a characterization of those functions in $\mathfrak{X}$ which vanish at x_∞ in the sense that $f^*(x_\infty) = 0$. It is easily seen, as a matter of fact, that the property in question is equivalent to the following property: corresponding to $\epsilon > 0$ there is a closed compact subset X_ϵ of X such that $|f(x)| < \epsilon$ for all x outside X. The totality of such functions is obviously a closed linear lattice ideal and a closed ring ideal in $\mathfrak{X}$, as we see by directly applying the results of §4 to $\mathfrak{X}^*$. We shall designate this class of functions as $\mathfrak{X}_\infty$. One of the most useful and typical approximation theorems for a locally compact, but not compact, space X is then stated as follows.

Theorem 12: *Let X be a locally compact, but not compact, space; and let $\mathfrak{X}$ and $\mathfrak{X}_\infty$ have the significance indicated above. If $\mathfrak{X}_0$ is any subfamily of $\mathfrak{X}_\infty$ which is a separating family for X and which contains for any x_0 in X a function f_0 such that $f_0(x_0) \neq 0$, then any function in $\mathfrak{X}_\infty$ can be uniformly approximated by linear lattice polynomials or by ordinary polynomials in members of $\mathfrak{X}_0$; and any function in $\mathfrak{X}$ can be similarly approximated by functions which result from the addition of a fixed constant function to such polynomials.*

Proof: If we look at $\mathfrak{X}^*$ and $\mathfrak{X}^*_\infty$, we see that $\mathfrak{X}^*_0$ is a separating family for X^* and that every function f^* in it vanishes at x_∞, $f^*(x_\infty) = 0$. By Theorems 3 and 7 we see that every function in $\mathfrak{X}^*_\infty$ can be uniformly approximated by linear-lattice polynomials in members of $\mathfrak{X}^*_0$. Similarly by Corollary 2 to Theorem 5 we obtain the corresponding statement for ordinary polynomials. If these results are now put in terms of $\mathfrak{X}$, $\mathfrak{X}_\infty$, and $\mathfrak{X}_0$, we obtain the present theorem. If f is in $\mathfrak{X}$, then we see that the function g defined by putting $g(x) = f(x) - f^*(x_\infty)$ for every x is in $\mathfrak{X}_\infty$ since $g^*(x_\infty) = f^*(x_\infty) - f^*(x_\infty) = 0$. Since g can be uniformly approximated by linear-lattice polynomials or ordinary polynomials in members of $\mathfrak{X}_0$, we conclude that f can be approximated by functions obtained by adding a constant function (everywhere equal to $f^*(x_\infty)$) to such a polynomial.

A useful special instance of this theorem may be phrased as follows.

Corollary 1: *Let X be a locally compact, but not compact, Hausdorff space; and let $\mathfrak{X}_0$ be a family of real continuous functions on X with the following properties: each function in $\mathfrak{X}_0$ vanishes outside a corresponding compact subset of X; corresponding to any point x_0 in X and any open set U which contains x_0 is a function in $\mathfrak{X}_0$ vanishing outside U and assuming at x_0 a value different from* 0. *Then $\mathfrak{X}_0$ has all the properties listed in the theorem above.*

Proof: Since any compact subset of X is now necessarily closed, it is evident that $\mathfrak{X}_0$ is part of $\mathfrak{X}_\infty$. It is easily verified that $\mathfrak{X}_0$ is a separating family for X, since when x_0 and x_0' are distinct points in X there is an open set U containing x_0 but not x_0' and hence a function f_0 in $\mathfrak{X}_0$ vanishing outside U and not vanishing at x_0, whence $f_0(x_0) \neq f_0(x_0') = 0$. We now see that $\mathfrak{X}_0$ satisfies the requirements laid down in Theorem 12, and the corollary is established.

For the case of complex-valued functions all the preceding remarks can evidently be repeated almost verbatim; the only essential changes which have to be made are to suppress all references to lattice-properties and to require of $\mathfrak{X}_0$ in Theorem 12 and the corollary that it contain $\bar{f}$ along with f. No further comment on this case would seem to be necessary.

7. *The Lebesgue-Urysohn Extension Theorem.* As a first application of the approximation theorems developed in the preceding sections we shall discuss a variant of the celebrated and important Lebesgue-Urysohn extension theorem, which asserts that corresponding to any continuous real function defined on a closed subset X_0 of a normal space X there exists a continuous real function defined on the entire space X and agreeing throughout X_0 with the given function. Since the known proofs of this theorem are quite simple (see, for example, Alexandroff-Hopf [5]), the present discussion is chiefly of interest in the realm of systematics. It would take us too far afield, in any event, to prove the extension theorem in its full generality, since the method to be applied consists in first compactifying the normal space X and then applying the results which we shall establish below. We shall therefore postpone the presentation of the full proof to some other occasion. Here we shall consider the case where X_0 is compact and X arbitrary, obtaining a result which in some respects actually goes beyond that summarized above as the Lebesgue-Urysohn extension theorem. We confine ourselves to the case of real functions, since the complex case is an essentially trivial consequence of it. Accordingly the theorem to be proved here can be stated as follows.

Theorem 13. *Let X be an arbitrary topological space, X_0 a compact subset of X, f_0 a continuous function defined on X_0, and $\mathfrak{X}_0$ a family of continuous real functions defined on X with the following properties: $\mathfrak{X}_0$ is closed under the linear lattice [linear ring] operations; $\mathfrak{X}_0$ is closed under uniform passage to the limit; $\mathfrak{X}_0$ contains all constant functions. In the case of the linear ring operations, let $\mathfrak{X}_0$ have further the property (verified automatically in the linear lattice case) that whenever it contains a function f it also contains a function g which coincides with f on X_0 and has the same bounds on X as f does on X_0. Then for f_0 to be extensible in $\mathfrak{X}_0$, in the sense that there exists a function f in $\mathfrak{X}_0$ agreeing with f_0 throughout X_0, it is necessary and sufficient that f_0 satisfy every linear condition of the form $f(x_0) = f(y_0)$, $x_0 \neq y_0$, $x_0 \in X_0$, $y_0 \in X_0$, which is satisfied by every member of $\mathfrak{X}_0$. The extension f can be so chosen as to have the same bounds on X as f_0 has on X_0.*

Proof: We first remark that the application of the linear lattice [linear ring] operations to extensible functions defined on X_0 produces extensible functions, since $\mathfrak{X}_0$ is assumed to be a linear lattice [linear ring]. We then observe that the uniform limit on X_0 of extensible functions is extensible. Indeed, let a sequence of functions g_n in $\mathfrak{X}_0$ converge uniformly on X_0 to a limit function g_0 (defined on X_0). We may suppose (by thinning out the originally given sequence, should that be necessary) that $|g_{n+1}(x) - g_n(x)| \leq 2^{-n}$ for every x in X_0. We then take h_n to be an extension of $g_{n+1} - g_n$ such that $|h_n| \leq 2^{-n}$. This is always possible, since we can choose h_n as a member of $\mathfrak{X}_0$ which has on X the same bounds as does $g_{n+1} - g_n$ on X_0; or, in the linear lattice case, we can simply take h_n so that

$$h_n(x) = \max(\min(g_{n+1}(x) - g_n(x),\ 2^{-n}),\ -2^{-n}).$$

Now the series $g_1 + \sum_{n=1}^{\infty} h_n$ converges uniformly to a continuous real function g in $\mathfrak{X}_0$ which agrees on X_0 with g_0. The family of all extensible functions on X_0 thus constitutes a closed linear sublattice [linear subring] of the linear lattice [linear ring] of all continuous real functions on X_0, and includes the constant functions on X_0. Theorem 3 and its Corollary in the lattice case, and Theorem 5 and its Corollary 1 in the ring case, show that the family of all extensible functions on X_0 coincides with the family of all those continuous functions on X_0 which satisfy every linear condition of the form $f(x_0) = f(y_0)$, $x_0 \not\leq y_0$, $x_0 \in X_0$, $y_0 \in Y_0$, which is satisfied by every function in $\mathfrak{X}_0$. The final statement of the theorem is obvious. In the lattice case, if α and β are the greatest lower and least upper bounds of f_0 on X_0 (both finite because X_0 is compact) and if f is any extension of f_0 in $\mathfrak{X}_0$, then f can be replaced by g where $g(x) = \max(\min f(x),\ \beta),\ \alpha)$.

The most interesting case of Theorem 13 is that in which $\mathfrak{X}_0$ is taken to be the family of all continuous functions on X. For this case we may state Theorem 13 in the following form.

Corollary 1. *In order that a continuous real function f_0 defined on a compact subset X_0 of a topological space X have a continuous extension defined over X, it is necessary and sufficient that f_0 satisfy every condition of the form $f(x_0) = f(y_0)$, $x_0 \not\leq y_0$, $x_0 \in X_0$, $y_0 \in Y_0$ which is satisfied by all continuous functions on X — in other words, that f_0 be constant on every subset of X_0 where all the functions continuous on X are constant by virtue of the topological structure of X. In particular, this condition is superfluous* *when the continuous functions on X constitute a separating family for X_0: every real function continuous on X_0 then has a continuous extension defined on X. If a function f_0 on X_0 has a continuous extension on X then it has such an extension with the same bounds on X as f_0 has on X_0.*

8. *The Theorem of Dieudonné.* A second interesting and useful application, still in the field of general topology, can now be made to a situation first adequately discussed by Dieudonné. Here we must presuppose the rudiments of the theory of the cartesian product of topological spaces. It is convenient to think of the product of the spaces X_α (where α runs over a fixed index-set A) as a coördinate space, each point x being specified by its coordinates x_α in the respective factor-spaces X_α. Now if α is a fixed index and f_α is a continuous real function defined on the factor space X_α, we can define a continuous real function f on the product space by putting $f(x) = f_\alpha(x_\alpha)$ where x_α is the coördinate of x corresponding to the index α. Such a function f will be called here a function of one variable — specifically, the function of one variable associated with f_α. These simple preliminaries enable us to state our main result as follows.

Theorem 14: *(Dieudonné, [6]). If X is the cartesian product of compact spaces X_α, $\alpha \in A$, then every continuous real function on X can be uniformly approximated by finite sums of finite products of continuous functions of one variable on X.*

Proof: The cartesian product of compact spaces is known to be compact. Let now $\mathfrak{X}_0$ be the totality of those functions expressible as finite sums of finite products of continuous functions of one variable on X. Since the sums, products, and constant multiples of functions in $\mathfrak{X}_0$ are obviously also in $\mathfrak{X}_0$, we see that $\mathfrak{X}_0$ is a linear subring of the ring $\mathfrak{X}$ of all continuous real functions on X. Obviously $\mathfrak{X}_0$ contains all the constant functions. In order to be able to apply the results of §3 we therefore have to determine what linear relations of the form $f(x) = f(x')$ where $x \neq x'$ are satisfied by every function f in $\mathfrak{X}_0$. Since we may take f here as the function of one variable associated with an arbitrary continuous real function f_α on X_α we must evidently have $f_\alpha(x_\alpha) = f_\alpha(x_\alpha')$ for all f_α. Conversely, if x and x' are points such that $f_\alpha(x_\alpha) = f_\alpha(x_\alpha')$ for every continuous real function f_α on X_α and for all α, then it is evident that $f(x) = f(x')$ for every f in $\mathfrak{X}_0$. Hence we see that any function in $\mathfrak{X}$ which satisfies all the linear relations of the above type can be uniformly approximated by functions in $\mathfrak{X}_0$, by virtue of Theorem 5. We can therefore complete our proof by showing that any continuous real function on X satisfies all these conditions. — First let us consider two points x and x' such that for some fixed index β we have $x_\beta \neq x_\beta'$ while $x_\alpha = x_\alpha'$ for all $\alpha \neq \beta$. From any function f on X we can obtain a function f_β on X_β by putting $f_\beta(y_\beta) = f(y)$ where y_β is arbitrary and $y_\alpha = x_\alpha = x_\alpha'$ for $\alpha \neq \beta$; and f_β is continuous when f is. Hence we see that if x_β and x_β' are points such that $f_\beta(x_\beta) = f_\beta(x_\beta')$ for every continuous real function f_β on X_β, then $f(x) = f_\beta(x_\beta) = f_\beta(x_\beta') = f(x')$. Let us suppose that we have generalized the result just established and have proved that, when x and x' are two points such that for fixed indices $\beta_1, \ldots, \beta_n$ we have $f_{\beta_k}(x_{\beta_k}) = f_{\beta_k}(x'_{\beta_k})$ for every continuous real function f_{β_k} on X_{β_k}, $k = 1, \ldots, n$ while $x_\alpha = x_\alpha'$ for every α other than $\beta_1, \ldots, \beta_n$, then $f(x) = f(x')$ for every f in $\mathfrak{X}$. We can then establish the corresponding result for points differing in at most $n + 1$ coördinates. In fact let x and x' be given so that $f_{\beta_k}(x_{\beta_k}) = f_{\beta_k}(x'_{\beta_k})$ as above for $k = 1,\ldots,n+1$ while $x_\alpha = x_\alpha'$ for all α other than $\beta_1,\ldots,\beta_{n+1}$. We define a point x'' by putting $x''_{\beta_k} = x_{\beta_k}$ for $k = 1,\ldots,n$, $x''_{\beta_{n+1}} = x'_{\beta_{n+1}}$ for $k = n+1$, and $x''_\alpha = x_\alpha = x'_\alpha$ for all α other than $\beta_1,\ldots,\beta_{n+1}$. It is obvious then that $f(x) = f(x'')$ for every f in $\mathfrak{X}$ by the result explicitly proved above. On the other hand the assumption we have made implies that $f(x') = f(x'')$ for every f in $\mathfrak{X}$. Hence we conclude that $f(x) = f(x')$ for every f in $\mathfrak{X}$ as we wished to show. By induction, therefore, we conclude that if two points x and x' differ only in respect to their coördinates for a finite number of indices for each of which $f_\beta(x_\beta) = f(x_\beta')$ for every continuous real function f_β on X_β, then $f(x) = f(x')$ for all f in $\mathfrak{X}$. — Finally let us suppose that f is in $\mathfrak{X}$ and that x and x' are points such that for every α and every continuous real function f_α on X_α the relation $f_\alpha(x_\alpha) = f_\alpha(x_\alpha')$ holds. Then we can determine for any positive ϵ a neighborhood U_ϵ of x such that $|f(x) - f(y)| < \epsilon$ for every y in U_ϵ, by virtue of the continuity of f. By the way in which the topology of X is defined, we can now determine a point x'' which is in U_ϵ and which nevertheless differs from x' only in respect to a finite number of coördinates. This end can indeed be achieved by designating appropriate indices $\beta_1,\ldots,\beta_n$ and putting $x''_{\beta_k} = x_{\beta_k}$ for $k = 1,\ldots,n$ while $x''_\alpha = x'_\alpha$ for all other indices α. We then have $f(x'') = f(x')$ in accordance with our previous results and hence also $|f(x) - f(x')| = |f(x) - f(x'')| < \epsilon$. Since ϵ is arbitrary we conclude that $f(x) = f(x')$, whatever the function f in $\mathfrak{X}$.

Useful variants of this theorem can be obtained by considering the case where some of the factor spaces are locally compact but not compact. As they would involve us in more extensive topological discussions than seem desirable here, we shall leave the matter at this point. — It will be useful, perhaps, to recall a classical application of the theorem just proved: in the theory of integral equations a standard procedure is to replace the kernel $K(x,y)$, assumed to be a continuous function of its arguments on the square $a \le x \le b$, $a \le y \le b$, by a uniformly good approximant of the form $K'(x,y) = F_1(x)G_1(y) + \cdots + F_n(x)G_n(y)$ where $F_1,\cdots,F_n$ and $G_1,\cdots,G_n$ are continuous functions on the interval $[a,b]$. Since the square is the cartesian product of the intervals $a\le x\le b$, $a\le y\le b$, the theorem of Dieudonné gives a direct justification for this device.

9. *The Weierstrass Approximation Theorem.* In the present section we propose to derive from the results of §3 a demonstration of the classical Weierstrass approximation theorem. In spite of the fact that we shall give a comparatively broad version of the theorem, everything we shall have to say is merely a direct specialization of previously established results to the case at hand. The steps of the general development which would have to be retained in a direct independent proof of the Weierstrass theorem will be indicated after the derivation of the theorem from §3 has been presented.

Theorem 15. *(Weierstrass, [7]). Let X be an arbitrary bounded closed subset of n-dimensional cartesian space, the coördinates of a general point being $x_1,\ldots,x_n$. Any continuous real function f defined on X can be uniformly approximated on X by polynomials in the variables $x_1,\ldots,x_n$. In case X contains the origin $x = (0,\ldots,0)$, the function f can be uniformly approximated by polynomials vanishing at the origin if and only if f itself vanishes at the origin. Otherwise f can be uniformly approximated by such polynomials without qualification.*

Proof: The functions $f_1, \ldots, f_n$, where $f_k(x) = x_k$, are continuous real functions of x. They constitute a separating family $\mathfrak{X}_0$ for X since $x = x'$ if and only if $x_k = f_k(x) = f_k(x') = x'_k$ for $k = 1, \ldots, n$. When X does not contain the origin, then we cannot have $f_1(x) = f_n(x) = 0$ for any x in X; but when X contains the origin we obtain $f_1(x) = \ldots = f_n(x) = 0$ by taking x as the origin. Since X is bounded and closed it is compact. Accordingly, the results of §3, particularly those stated in Corollary 2 to Theorem 5, show that any continuous real function f on X can be uniformly approximated by functions of the form (where $\alpha_1, \ldots, \alpha_n$ are positive integers or 0)

$$p(x) = \sum_{1 \leq \alpha_1 + \cdots + \alpha_n \leq N} C_{\alpha_1 \ldots \alpha_n} (f_1(x))^{\alpha_1} \cdots (f_n(x))^{\alpha_n}$$
$$= \sum_{1 \leq \alpha_1 + \cdots + \alpha_n \leq N} C_{\alpha_1 \ldots \alpha_n} x_1^{\alpha_1} \cdots x_n^{\alpha_n},$$

with the proviso that when X contains the origin f must vanish there. If the constant function everywhere equal to 1 is adjoined to the family $\mathfrak{X}_0$, we see that f can be uniformly approximated on X by functions of the form

$$p(x) = \sum_{0 \leq \alpha_1 + \cdots + \alpha_n \leq N} C_{\alpha_1 \ldots \alpha_n} x_1^{\alpha_1} \cdots x_n^{\alpha_n}$$

This completes the proof.

If one wishes to give a direct proof of the Weierstrass approximation theorem by the present methods, the following procedure is available. It is best to consider first the family $\mathfrak{X}_0$ of all homogeneous linear functions, l, where $l(x) = c_1x_1 + \cdots + c_nx_n$, noting that for any given continuous real function f on X (provided f vanishes at the origin when this point is in X) a function l can be found so that $l(x) = f(x)$, $l(y) = f(y)$ at arbitrarily prescribed points x, y in X. Considerations like those used in the proof of Theorem 1 show that such a function f can be uniformly approximated by lattice combinations of functions in $\mathfrak{X}_0$. Theorem 4 has to be established exactly as in §3. It can then be used to convert approximation by lattice combinations into approximation by linear-ring combinations of functions in $\mathfrak{X}_0$, just as was done in the proof of Theorem 5. The main part of the proof is thereby completed. Remarks similar to those above have to be added concerning the adjunction of constant functions to $\mathfrak{X}_0$.

10. *Trigonometric Approximation.* A surprisingly direct and simple application of Theorem 5 yields the fundamental theorem on trigonometric approximation in the real domain, reading as follows.

Theorem 16. *Let f be an arbitrary continuous real function of the real variable θ, $0 \leq \theta \leq 2\pi$, subject to the periodicity condition $f(0) = f(2\pi)$. Then f can be uniformly approximated on its domain of definition by trigonometric polynomials — that is, by functions of the form*

$$p(\theta) = \alpha_0/2 + \sum_{n=1}^{N} (\alpha_n \cos n\theta + b_n \sin n\theta).$$

Proof: It is convenient to make the application of Theorem 5 directly to the case of the unit circle X (given by the equation $x_1^2 + x_2^2 = 1$) in the cartesian plane, thus obtaining a special instance of the Weierstrass approximation theorem as stated in Theorem 15. If we rephrase this special case in terms of the central angle θ corresponding to a general point x on the circle X, we see that the functions continuous on the circle are the functions continuous and periodic in θ

and that the polynomials in x_1 and x_2 (in terms of which such functions can be uniformly approximated) are functions of the form

$p(\theta) = \sum_{1 \le m+n \le N} c_{mn} \cos^m\theta \sin^n\theta$ (since $x_1 = \cos\theta$ and $x_2 = \sin\theta$). The addition formulas for the trigonometric functions yield the relations

$$2\cos m\theta \cos n\theta = \cos(m+n)\theta + \cos(m-n)\theta$$
$$2\cos m\theta \sin n\theta = \sin(m+n)\theta + \sin(m-n)\theta$$
$$2\sin m\theta \sin n\theta = -\cos(m+n)\theta + \cos(m-n)\theta$$

which enable us to extablish by a recursive argument that $\cos^m\theta \sin^n\theta$ (and hence also every function of the form $p(\theta)$ described above) is a trigonometric polynomial in the sense required by the statement of the theorem. Indeed, we have only to note that, if $\cos^m\theta \sin^n\theta$ is such a polynomial, then $\cos^{m+1}\theta \sin^n\theta$ and $\cos^m\theta \sin^{n+1}\theta$ are also trigonometric polynomials by virtue of the indicated relations.

The complex form of the trigonometric approximation theorem can be deduced even more readily from Theorem 10, Corollary 2; it can be stated as follows.

Theorem 17. *If f is a continuous complex function of the real variable θ, $0 \le \theta \le 2$, subject to the periodicity condition $f(0) = f(2\pi)$ — then f can be uniformly approximated by functions p of the form $p(\theta) = \sum_{n=-N}^{n=+N} C_n e^{ni\theta}$, where the constants C_n are complex numbers.*

Proof: The functions considered will be treated as continuous functions on the unit circle X, θ being the central angle as in the discussion of the preceding theorem. Since $e^{mi\theta} e^{ni\theta} = e^{(m+n)i\theta}$ and $\overline{e^{ni\theta}} = e^{-ni\theta}$, it is clear that the family $\mathfrak{X}_0$ of all functions p of the form $p(\theta) = \sum_{n=-N}^{n=N} C_n e^{ni\theta}$ is a linear subring of the family $\mathfrak{X}$ of all continuous complex functions on X, the function $\overline{p}$ being in $\mathfrak{X}_0$ whenever p is. It is also evident that the functions $e^{ni\theta}$ satisfy no linear relation of the form $e^{ni\theta_0} = 0$ or of the form $e^{ni\theta_1} = e^{ni\theta_2}$, $\theta_1 \ne \theta_2$, where $0 \le \theta_1 < 2\pi$, $0 \le \theta_2 < 2\pi$. Consequently, every function in $\mathfrak{X}$ can be uniformly approximated by functions in $\mathfrak{X}_0$, in accordance with Corollary 2 to Theorem 10. This completes the proof.

11. *Approximation by Laguerre Functions.* An important problem of analysis concerns the approximation of real functions continuous on the half-infinite interval $0 \le x < +\infty$ by linear combinations of the functions $e^{-\alpha x} x^n$, $n = 0,1,2,\ldots$, where α is a fixed positive number. In order to obtain a solution to this problem within the scope of the present discussion we need a lemma concerning the exponential function.

Lemma 1. *On the interval $0 \le x < \infty$ the function $e^{-n\alpha x}$, where n is a positive integer and $\alpha < 0$, can be uniformly approximated by functions of the form $e^{-\alpha x} p(x)$ where p is a polynomial.*

Proof. We may suppose without loss of generality that $\alpha = 1$. Indeed, if for $\epsilon < 0$ we have found a polynomial $q(x)$ such that $|e^{-nx} - e^{-x} q(x)| < \epsilon$ for $0 \le x < \infty$, we can replace x by αx and $q(x)$ by the polynomial $p(x) = q(\alpha x)$, obtaining $|e^{-n\alpha x} - e^{-\alpha x} p(x)| < \epsilon$ for $0 \le x < \infty$. We shall now proceed recursively. When $n = 1$ there is nothing to prove, since e^{-x} is already of the form specified for its approximants. When $n = 2$, we obtain the desired result by estimating the magnitude of the function f where $f(x) = e^{-2x} - e^{-x} \sum_{k=0}^{N} \frac{(-x)^k}{k!}$. Since f is continuous and has the properties $f(0) = 0$ and $\lim_{x\to\infty} f(x) = 0$, it has an extremum on the interval $0 < x < \infty$. If such an extremum occurs at $x = x_0$, then $0 = f'(x_0) = -2e^{-2x_0} + 2e^{-x_0} \sum_{k=0}^{N-1} \frac{(-x_0)^k}{k!} + e^{-x_0} \frac{(-x_0)^N}{N!}$ so that $f(x_0) = \frac{1}{2} e^{-x_0} \frac{(-x_0)^N}{N!}$. Consequently we have

$\sup_{0\le x<\infty} |f(x)| \le \frac{1}{2N!} \sup_{0\le x<\infty} e^{-x} x^N$. For $N \ge 1$ the function g defined by putting $g(x) = e^{-x} x^N$ is continuous and non-negative on the interval $0 \le x < \infty$ and has the properties $g(0) = 0$, $\lim_{x\to\infty} g(x) = 0$. It therefore has a maximum on the interval $0 < x < \infty$, occurring at the only solution there of the equation $0 = g'(x) = e^{-x} (N - x) x^{N-1}$ — that is, at $x = N$. Accordingly we have $0 \le g(x) \le e^{-N} N^N$. Applying so much of Stirling's formula as is necessary to show that $N! \le e^{-N} N^{N+\frac{1}{2}}/2K$ for a suitable constant $K > 0$ we obtain the inequality $|f(x)| \le KN^{-\frac{1}{2}}$. The case $n = 2$ is thereby settled. If we have proved the lemma for any particular positive integer n, we can discuss the approximation of $e^{-(n+1)x}$ in the following manner. If $\epsilon > 0$ is given, we first use what is known about the approximation of $e^{-2\alpha x}$, taking $2\alpha = n+1$, so as to obtain a polynomial q such that $|e^{-(n+1)x} - e^{-nx/2 - x/2} q(x)| < \frac{1}{2}\epsilon$. The function $e^{-x/2} |q(x)|$ is bounded on the interval $0 \le x < \infty$. If its least upper bound is A, we use what is known about the approximation of $e^{-n\alpha x}$, taking $\alpha = \frac{1}{2}$, so as to obtain a polynomial r such that $|e^{-nx/2} - e^{-x/2} r(x)| < \epsilon/2A$. It follows that $|e^{-nx/2 - x/2} q(x) - e^{-x} q(x) r(x)| < e^{-x/2} |q(x)| \epsilon/2A \le \frac{1}{2}\epsilon$ and $|e^{-(n+1)x} - e^{-x} p(x)| < \epsilon$, where $p = qr$. Mathematical induction therefore serves to complete the proof of the lemma.

A direct application of Theorem 12 together with the lemma just proved yields the main approximation theorem of this section.

Theorem 18. *Any continuous real function f which is defined on the interval $0 \le x < \infty$ and vanishes at infinity in the sense that $\lim_{x\to\infty} f(x) = 0$ can be uniformly approximated by functions of the form $e^{-\alpha x} p(x)$ where $p(x)$ is a polynomial.*

Proof: We let X be the interval $0 \le x < \infty$. As a topological space, X is locally compact. The function $e^{-\alpha x}$ is in $\mathfrak{X}_\infty$ since $\lim_{x\to 0} e^{-\alpha x} = 0$. We now let $\mathfrak{X}_0$ consist of this function alone. It is obvious from the monotonicity of the exponential function that $\mathfrak{X}_0$ is a separating family for X. Moreover there is no $x \ge 0$ for which $e^{-\alpha x} = 0$. By hypotheses the function f to be approximated is in $\mathfrak{X}_\infty$. Theorem 12 thus shows that f can be uniformly approximated on X by functions of the form $\sum_{n=1}^{N} C_n e^{-n\alpha x}$. Lemma 1 then yields the present theorem.

A variant of the above proof can be based on a direct appeal to the Weierstrass approximation theorem. We introduce a new variable $\xi = e^{-\alpha x}$, $0 < \xi \leq 1$. The function ϕ defined by putting $\phi(0) = 0$, $\phi(\xi) = f(x) = f(\frac{-1}{\alpha} \log \xi)$ is continuous on the interval $0 \leq \xi \leq 1$. Hence ϕ can be uniformly approximated by polynomials $\sum_{n=1}^{N} C_n \xi^n$; and f can be uniformly approximated by functions of the form $\sum_{n=1}^{N} C_n e^{-n\alpha\xi}$. Lemma 1 is then used to complete the proof.

It is of some interest to apply the approximation theorem just proved to derive results concerning approximation in the mean. The classical theorem on this subject reads as follows.

Theorem 19. *The functions of the form $e^{-\alpha x}p(x)$ where $p(x)$ is a polynomial are dense in the function-space $\mathfrak{L}_r(0,\infty)$, $r \geq 1$.*

Proof. Here $\mathfrak{L}_r(0,\infty)$ is the class of all real Lebesgue-measurable functions f on the interval $0 \leq x < \infty$ for which the Lebesgue integral $\int_0^\infty |f(x)|^r dx$ exists. The expression $\left[\int_0^\infty |f(x)|^r dx\right]^{1/r}$ is taken as the norm of f, and $\mathfrak{L}_r(0,\infty)$ then becomes a complete normed linear vector space. It is obvious that every function of the form $e^{-\alpha x}p(x)$, where p is a polynomial, is a member of $\mathfrak{L}_r(0,\infty)$. We wish to prove that if $\epsilon > 0$ and if f is in $\mathfrak{L}_r(0,\infty)$ then there exists a function of this special form for which $\left[\int_0^\infty |f(x) - e^{-\alpha x}p(x)|^r dx\right]^{1/r} < \epsilon$. It is well known that there exists a function g, continuous on the interval $0 \leq x < \infty$ and vanishing outside some bounded interval, for which $\left[\int_0^\infty |f(x) - g(x)|^r dx\right]^{1/r} < \frac{1}{2}\epsilon$. We therefore need only show that it is possible to find a function of the indicated special form for which $\left[\int_0^\infty |g(x) - e^{-\alpha x}p(x)|^r\right]^{1/r} < \frac{1}{2}\epsilon$. The function $e^{\frac{1}{2}\alpha x}g(x)$ is continuous on $0 \leq x < \infty$ and vanishes outside some bounded interval. Hence we can find a polynomial $p(x)$ such that $|e^{\frac{1}{2}\alpha x}g(x) - e^{-\frac{1}{2}\alpha x}p(x)| < \eta$ where $\eta < \frac{1}{2}\left(\frac{r\alpha}{2}\right)^{1/r}\epsilon$. Accordingly we obtain the relations

$$\left[\int_0^\infty |g(x) - e^{-\alpha x}p(x)|^r\right]^{1/r} = \left[\int_0^\infty |e^{\frac{1}{2}\alpha x}g(x) - e^{-\frac{1}{2}\alpha x}p(x)|^r e^{-\frac{r\alpha x}{2}} dx\right]^{1/r} \leq$$

$$\eta\left[\int_0^\infty e^{-\frac{r\alpha x}{2}} dx\right]^{1/r} = \left[\frac{2}{r\alpha}\right]^{1/r}\eta < \tfrac{1}{2}\epsilon.$$

The proof is thereby completed.

A useful related theorem is the following.

Theorem 20. *If f is in $\mathfrak{L}_r(0,\infty)$, $r \geq 1$, or if f is a bounded Lebesgue-measurable function, then the integrals $\int_0^\infty f(x)e^{-\alpha x}x^n dx$ exist for $\alpha > 0$ and $n = 0,1,2,\cdots$. If these integrals vanish for fixed α and $n = 0,1,2,\cdots$, then $f(x)$ vanishes almost everywhere.*

Proof. When f is in $\mathfrak{L}_1(0,\infty)$, the fact that the functions $e^{-\alpha x}x^n$ are bounded yields the existence of the integrals in question. When f is in $\mathfrak{L}_r(0,\infty)$ for $r > 1$, the fact that the functions $e^{-\alpha x}x^n$ are in $\mathfrak{L}_{r'}(0,\infty)$ where $1/r + 1/r' = 1$, yields a like result in the standard way. Finally when f is bounded and Lebesgue-measurable, it is the fact that the functions $e^{-\alpha x}x^n$ are in $\mathfrak{L}_1(0,\infty)$ which yields the desired result. Using this result we see that in every case the function $g(x) = e^{-\frac{1}{4}\alpha x}\int_0^x f(t)e^{-\frac{1}{4}\alpha t}dt$ is a continuous function with the property that $|g(x)| \leq Ke^{-\frac{1}{4}\alpha x}$ for some constant K. Thus g is in $\mathfrak{L}_2(0,\infty)$. Moreover an integration by parts shows that

$$\begin{aligned}\int_0^\infty g(x)e^{-\frac{1}{2}\alpha x}x^n dx &= \int_0^\infty \left(\int_0^x f(t)e^{-\frac{1}{4}\alpha t}dt\right)(e^{-\frac{3}{4}\alpha x}x^n)dx \\ &= \int_0^\infty (f(x)e^{-\frac{1}{4}\alpha x})\left(\int_0^x e^{-\frac{3}{4}\alpha t}t^n dt\right)dx \\ &= \int_0^\infty f(x)e^{-\alpha x}p(x)dx = 0\end{aligned}$$

since $\int_0^x e^{-\frac{3}{4}\alpha t}t^n dt = e^{-\frac{3}{4}\alpha x}p(x)$ where $p(x)$ is a polynomial and since it is assumed that $\int_0^\infty f(x)e^{-\alpha x}x^n dx = 0$ for $n = 0,1,2,\cdots$. We have thus reduced the proof of the theorem to the special case where the given function is in $\mathfrak{L}_2(0,\infty)$. Choosing a polynomial $p(x)$ such that $\left[\int_0^\infty |g(x) - e^{-\frac{1}{2}\alpha x}p(x)|^2 dx\right]^{1/2} < \epsilon$ and noting that $\int_0^\infty g(x)e^{-\frac{1}{2}\alpha x}p(x)dx = 0$, we have $\int_0^\infty |g(x)|^2 dx =$ $\int_0^\infty \left[g(x) - e^{-\frac{1}{2}\alpha x}p(x)\right] g(x)dx \leq \left[\int_0^\infty |g(x) - e^{-\frac{1}{2}\alpha x}p(x)|^2 dx\right]^{1/2} \left[\int_0^\infty |g(x)|^2 dx\right]^{1/2}$ by Schwarz's inequality; and we therefore have $\left[\int_0^\infty |g(x)|^2 dx\right]^{1/2} < \epsilon$. It follows that $\int_0^\infty |g(x)|^2 dx = 0$ and that $g(x)$, being continuous, vanishes identically. The relation $\int_0^x f(t)e^{-\frac{1}{4}\alpha t}dt = 0$ is thus established. From this it follows that $f(x)e^{-\frac{1}{4}\alpha x} = 0$ and $f(x) = 0$, almost everywhere.

In the three theorems proved in this section, it is obvious that the hypotheses concerning the function f can be altered to allow f to be complex, without changing the conclusions.

12. *Approximation by Hermite Functions.* The methods of the preceding section can be applied with little modification to yield comparable results concerning uniform approximation by linear combinations of the functions $e^{-\alpha^2 x^2}x^n$ on the full infinite interval $-\infty < x < +\infty$. Using Theorem 12 and Lemma 1, we immediately obtain the chief result.

Theorem 21. *Any continuous real function f which is defined on the interval $-\infty < x < +\infty$ and which vanishes at infinity in the sense that $\lim_{x\to-\infty} f(x) = \lim_{x\to+\infty} f(x) = 0$ can be uniformly approximated by functions of the form $e^{-\alpha^2x^2}p(x)$*

where $p(x)$ is a polynomial.

Proof: We let X be the interval $-\infty < x < +\infty$. As a topological space, X is locally compact. The functions f, $e^{-\alpha^2x^2}$, and $e^{-\alpha^2x^2}x$ are in $\mathfrak{X}_\infty$ since they all vanish at infinity in the sense indicated in the statement of the theorem. The family $\mathfrak{X}_0$ consisting of the two functions $e^{-\alpha^2x^2}$, $e^{-\alpha^2x^2}x$ is obviously a separating family for X; and moreover there is no x for which $e^{-\alpha^2x^2} = 0$. Theorem 12 thus shows that f can be uniformly approximated by functions of the form $\sum_{m=1}^{M}\sum_{n=1}^{N} c_{mn}e^{-m\alpha^2x^2}x^n$. To complete the discussion we make use of Lemma 1. Letting A be the maximum of the function $e^{-\frac{1}{2}\alpha^2x^2}|x|^n$, we find a polynomial q such that $|e^{-(2m-1)t}-e^{-t}q(t)| < \epsilon/A$ for $0 \le t < \infty$. We now put $t = \frac{1}{2}\alpha^2x^2$, multiply both sides of the inequality by $e^{-\frac{1}{2}\alpha^2x^2}|x|^n$, and write $p(x)$ for $x^nq(\frac{1}{2}\alpha^2x^2)$, obtaining

$$|e^{-m\alpha^2x^2}x^n - e^{-\alpha^2x^2}p(x)| < e^{-\frac{1}{2}\alpha^2x^2}|x|^n\,\epsilon/A \le \epsilon.$$

The theorem then follows.

A variant of this proof can be given by appropriate use of the Weierstrass approximation theorem, as stated in §9. We introduce a variable point ξ of the cartesian plane with the coordinates $x_1 = e^{-\alpha^2x^2}$, $x_2 = e^{-\alpha^2x^2}x$. The locus of this point ξ, with the origin adjoined, provides a bounded closed set χ in the plane. The function ϕ defined by putting $\phi(\xi) = f(x)$ when $\xi = (x_1, x_2) = (e^{-\alpha^2x^2}, e^{-\alpha^2x^2}x)$ and $\phi(\xi) = 0$ when ξ is the origin is a continuous function of ξ on X. The Weierstrass approximation theorem shows that ϕ can be uniformly approximated on X by a polynomial $\sum_{l=1}^{L}\sum_{n=1}^{N} c'_{ln}x_1^l x_2^n$. Hence f can be uniformly approximated on the interval $-\infty < x < +\infty$ by a function of the form $\sum_{l=1}^{L}\sum_{n=1}^{N} c'_{ln}e^{-(l+n)\alpha^2x^2}x^n = \sum_{m=1}^{M}\sum_{n=1}^{N} c_{mn}e^{-m\alpha^2x^2}x^n$, where $m = l+n$, $M = L+N$, and $c_{mn} = c'_{m-n,n}$. The remainder of the proof is identical with that given above.

By a method almost the same as that used in proving Theorem 19, we obtain the corresponding result for approximation in the mean by functions of the form $e^{-\alpha^2x^2}p(x)$ where p is a polynomial. The theorem is therefore stated without detailed proof.

Theorem 22. *The functions of the form $e^{-\alpha^2x^2}p(x)$, where $p(x)$ is a polynomial, are dense in the function-space $\mathfrak{L}_r(-\infty, +\infty)$, $r \ge 1$.*

By using the results of Theorem 20 we can give a simple proof of its analogue for the functions $e^{-\alpha^2x^2}x^n$.

Theorem 23. *If f is in $\mathfrak{L}_r(-\infty,+\infty)$, $r \ge 1$, or if f is a bounded Lebesgue-measurable function, then the integrals $\int_{-\infty}^{+\infty} f(x)e^{-\alpha^2x^2}x^ndx$ exist for $\alpha > 0$ and $n = 0,1,2,\cdots$. If these integrals vanish for fixed α and $n = 0,1,2,\cdots$, then $f(x)$ vanishes almost everywhere.*

Proof: Let f_1 and f_2 be the functions defined by putting $2f_1(x) = f(x) + f(-x)$, $2f_2(x) = f(x) - f(-x)$, so that f_1 is even, f_2 is odd, and $f = f_1 + f_2$. It is evident that f_1 and f_2 are in $\mathfrak{L}_r(-\infty,+\infty)$ when f is, and are bounded when f is. We easily see that

$$2\int_0^\infty f_2(x)e^{-a^2x^2}x^{2m+1}dx = \int_{-\infty}^{+\infty} f_2(x)e^{-a^2x^2}x^{2m+1}dx = \int_{-\infty}^{+\infty} f(x)e^{-a^2x^2}x^{2m+1}dx = 0.$$

The function f_2, considered on the half-infinite interval $0 \le x < \infty$, is in $\mathfrak{L}_r(0,\infty)$ or is bounded according as f_2 is in $\mathfrak{L}_r(-\infty,+\infty)$ or is bounded. We now make the change of variable $t = x^2$ and write $g(t) = f_2(x)e^{-\frac{1}{2}a^2x^2}$, obtaining

$$\int_0^\infty g(t)e^{-\frac{1}{2}a^2t}t^m dt = 2\int_0^\infty f_2(x)e^{-a^2x^2}x^{2m+1}dx = 0,$$

$$\int_0^\infty |g(t)|^r dt = 2\int_0^\infty |f_2(x)|^r e^{-\frac{r}{2}a^2x^2}x dx < +\infty$$

so that g is in $\mathfrak{L}_r(0,\infty)$ if f is and is bounded if f is. Thus Theorem 20 is applicable and yields the result that $g(x)$ vanishes almost everywhere. It follows that f is essentially an even function of x, the equation $f(-x) = f(x)$ being satisfied almost everywhere. It is easy to see that the function h defined by putting $h(x) = f(x)e^{-\frac{1}{2}a^2x^2}x$ is essentially odd; that h is in $\mathfrak{L}_r(-\infty,+\infty)$ when f is and is bounded when f is; and that $\int_{-\infty}^{+\infty} h(x)e^{-\frac{1}{2}a^2x^2}x^n dx = 0$ for $n = 0,1,2,\cdots$. By what has already been proved h must be essentially even. However, since it was given as an essentially odd function, it must vanish almost everywhere. Thus $f(x) = 0$ almost everywhere, as we wished to prove.

It is obvious that in the three theorems proved in this section the hypotheses concerning the function f can be altered so as to allow f to be complex without changing the conclusions.

13. *The Peter-Weyl Approximation Theorem.* As a final application of our results concerning approximation, we shall sketch briefly a proof of the theorem of Peter and Weyl concerning the approximation of functions on a compact topological group*. This theorem includes as a special case the classical theorems on trigonometric approximation given in §10, as is well-known.

Let X be a compact topological group. It is known that this group has a complete system of mutually inequivalent, irreducible continuous real (respectively, complex) matrix representations by finite orthogonal (respectively, unitary) matrices. Specifically a system of finite orthogonal (respectively, unitary) matrices $A^{(k)}(x)$, where x is in X and $k = 1,2,3,\cdots$, can be found with the following properties:

(1) the elements of $A^{(k)}(x)$ depend continuously on x; the relation $A^{(k)}(x)A^{(k)}(y) = A^{(k)}(xy)$ is satisfied; and $A^{(k)}(e)$, e being the identy element of X, is a unit matrix;

(2) the continuous representation of X given by $A^{(k)}(x)$ is irreducible;

(3) the continuous representations of X given by $A^{(k)}(x)$ and $A^{(k')}(x)$ are inequivalent when $k \neq k'$;

*The proof offered here was presented by the author to a seminar held at the University of Buenos Aires in 1943.

(4) any irreducible continuous real (respectively, complex) matrix representation of X is equivalent to the representation given by $\Lambda^{(k)}(x)$ for some k (necessarily unique);

(5) any two distinct elements x and y of the group X determine at least one k such that $\Lambda^{(k)}(x) \neq \Lambda^{(k)}(y)$.

The problem to be discussed is that of approximating a general continuous function on X in terms of the functions $\lambda_{ij}^{(k)}$ defined by taking $\lambda_{ij}^{(k)}(x)$ as the element standing in the i^{th} row and j^{th} column of the matrix $\Lambda^{(k)}(x)$. We shall therefore take $\mathfrak{X}$ to be the family of all continuous real (respectively, complex) functions on X, and $\mathfrak{X}_0$ as the family comprising all the functions $\lambda_{ij}^{(k)}$, $i = 1,\cdots,n_k$, $j = 1,\cdots,n_k$, $k = 1,2,3,\cdots$. By virtue of (1) we see that $\mathfrak{X}_0$ is part of $\mathfrak{X}$. By virtue of (5) we see that $\mathfrak{X}_0$ is a separating family for X. A trivial irreducible representation of X can be obtained by letting $\Lambda(x)$ be the matrix of one row and one column whose single element has the value 1. From (1) and (4) we must have $\Lambda(x) = \Lambda^{(k)}(x)$ for some k, necessarily unique. Hence $\mathfrak{X}_0$ contains the constant function which assumes the value 1 everywhere on X. In the complex case, it is evident that the matrix $\overline{\Lambda^{(k)}(x)}$ whose elements are the conjugates of the elements of $\Lambda^{(k)}(x)$ gives an irreducible unitary representation of X, by virtue of the relation $\overline{\Lambda^{(k)}(x)}\,\overline{\Lambda^{(k)}(y)} = \overline{\Lambda^{(k)}(xy)}$. Because of (4) this representation is equivalent to that given by the matrices $\Lambda^{(k')}(x)$ for some k, necessarily unique. Thus there is a non-singular constant matrix Λ such that $\overline{\Lambda^{(k)}(x)} = \Lambda^{-1}\Lambda^{(k')}(x)\Lambda$. Accordingly each of the functions $\overline{\lambda_{ij}^{(k)}}$ is a linear combination (with constant coefficients) of the functions $\lambda_{i'j'}^{(k')}$, $i' = 1,\cdots,n_k'$, $j' = 1,\cdots,n_k'$. The most important property of all is expressed by a similar statement — namely, that the product $\lambda_{ij}^{(k)}\lambda_{i'j'}^{(k')}$ is a finite linear combination with real (respectively, complex) constant coefficients of the functions $\lambda_{i''j''}^{(k'')}$. The proof of this assertion is obtained by considering the representation of X given by the Kronecker product $\Lambda^{(k)}(x) \times \Lambda^{(k')}(x)$ of the matrices $\Lambda^{(k)}(x)$ and $\Lambda^{(k')}(x)$. The Kronecker product in question is a matrix of $n_k^2 n_k^2$ elements whose rows and columns are labeled by the pairs (i,i'), (j,j') respectively, the element standing in the row labeled (i,i') and column labeled (j,j') being $\lambda_{ij}^{(k)}(x)\lambda_{i'j'}^{(k')}(x)$. By direct computation it is easy to verify that

$$\left[\Lambda^{(k)}(x) \times \Lambda^{(k')}(x)\right]\left[\Lambda^{(k)}(y) \times \Lambda^{(k')}(y)\right] =$$

$$\left[\Lambda^{(k)}(x)\Lambda^{(k)}(y)\right]\times\left[\Lambda^{(k')}(x)\Lambda^{(k')}(y)\right] = \Lambda^{(k)}(xy) \times \Lambda^{(k')}(xy).$$

Hence the Kronecker product provides a continuous representation of X. When this representation is resolved into its irreducible constituents, finite in number, each of the latter is equivalent in accordance with (4) to a representation $\Lambda^{(k'')}(x)$ for some k'', necessarily unique. This resolution corresponds to the determination of a matrix Λ of real (respectively, complex) constants such that the matrices $\Lambda\left[\Lambda^{(k)}(x) \times \Lambda^{(k')}(x)\right]\Lambda^{-1} = \Lambda(x)$ have the form indicated schematically as follows:

$$\begin{pmatrix} \Lambda^{(k_1'')}(x) & 0 & 0 & 0 \\ 0 & \Lambda^{(k_2'')}(x) & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \Lambda^{(k_n'')}(x) \end{pmatrix}$$

Here the blocks along the principal diagonal are occupied by various ones of the matrices $\Lambda^{(k'')}(x)$ (not necessarily distinct!) and all other blocks are filled with zeros. The fact that $\Lambda^{(k)}(x)\Lambda^{(k')}(x) = \Lambda^{-1}\Lambda(x)\Lambda$ leads at once to the conclusion that $\lambda_{ij}^{(k)}\lambda_{i'j'}^{(k')}$ is a real (respectively, complex) linear combination (with constant coefficients) of functions $\lambda_{i''j''}^{(k'')}$ corresponding to elements of the diagonal blocks in the matrix $\Lambda(x)$.

From the properties enumerated above we can now obtain the approximation theorem by direct application of Theorems 5 and 10.

Theorem 24 *(Peter-Weyl, [8]). Any continuous real (respectively, complex) function on the compact topological group X can be uniformly approximated on X by real (respectively, complex) linear combinations of the functions $\lambda_{ij}^{(k)}$ described above in terms of the irreducible representations of X.*

Proof. Let $\mathfrak{U}(\mathfrak{X}_0)$ be the family of all the real (respectively, complex) continuous functions on X which can be uniformly approximated by polynomials in members of $\mathfrak{X}_0$, the family of all the functions $\lambda_{ij}^{(k)}$. In view of the facts about the products $\lambda_{ij}^{(k)}\lambda_{i'j'}^{(k')}$, it is clear that $\mathfrak{U}(\mathfrak{X}_0)$ is also the family of all the real (respectively, complex) continuous functions on X which can be uniformly approximated by real (respectively, complex) linear combinations of the functions $\lambda_{ij}^{(k)}$. In the complex case, we know further that $\mathfrak{U}(\mathfrak{X}_0)$ contains $\bar{f}$ along with f. Since $\mathfrak{X}_0$ is a separating family for X and contains non-vanishing constant functions, Theorem 5 shows that $\mathfrak{U}(\mathfrak{X}_0) = \mathfrak{X}$ in the real case and Theorem 10 leads to the same result in the complex case.

14. *Linear Combinations of Prescribed Functions.* It would be natural to study, by way of further generalizing the results obtained here, the problem of approximation in terms of linear combinations of prescribed functions. In this domain, however, there are encountered some of the most difficult problems of analysis. For example, Wiener [9] has shown that general Tauberian theorems are intimately related to the problem of approximation in the mean by linear combinations of functions f_a obtained from a single function f in $\mathfrak{L}_1(-\infty, +\infty)$ by putting $f_a(x) = f(x - a)$. The conditions under which every function in $\mathfrak{L}_1(-\infty, +\infty)$ can be approximated in the mean (of order one) by such linear combinations were obtained by Wiener with the use of ingenious and powerful methods. Modern versions of his treatment have brought many simplifications, but still leave the impression that the results are among the deeper achievements of analysis.

In general, therefore, one cannot expect that the theory of this broader problem will assume so satisfactory a form as that which has been worked out when the lattice or the ring operations can be used to build approximants. The fact that we have been able in §§10-13 to apply our theory to obtain results concerning particular cases of the broader problem is simply due to the observation that under certain circumstances it is possible to approximate products of the prescribed functions by linear combinations of them. This observation leads to an application of the ring theorems given in §3 and §5, in the manner exemplified in §§10-13. Whenever a special theorem concerning the approximation of products by linear combinations can be established, the way is open for the employment of the same device.

REFERENCES

[1] Stone, M. H., Transactions of the American Mathematical Society, 41 (1937), pp. 375-481, especially pp. 453-481.

[2] Kakutani, Shizuo, Annals of Mathematics, (2) 42 (1941), pp. 994-1024, especially pp. 1004-1005.

[3] Šilov, G., Comptes Rendus (Doklady), Akademia Nauk, U.S.S.R., 22 (1939), pp. 7-10.

[4] Čech, E., Annals of Mathematics, (2) 38 (1937), pp. 823-844.

[5] Alexandroff, P., and Hopf, H., "Topologie I," (Berlin, 1935), pp. 73-78.

[6] Dieudonné, J., Comptes Rendus (Paris), 205 (1937), p. 593.

[7] Weierstrass, Karl, Mathematische Werke, Band 3, Abhandlungen III, pp. 1-37, especially p. 5 (= Sitzungsberichte, Kön. Preussischen Akademie der Wissenschaften, July 9 and July 30, 1885).

[8] Peter, F., and Weyl, H., Mathematische Annalen, 97 (1927), pp. 737-755, especially p. 753.

[9] Wiener, Norbert, "The Fourier Integral," (Cambridge, England, 1933) especially Chapter II, §14.

University of Chicago
Chicago 37, Illinois