THE GENERAL THEORY OF APPROXIMATION BY POLYNOMIALS AND TRIGONOMETRIC SUMS.

REPORT PRESENTED BEFORE THE AMERICAN MATHEMATICAL SOCIETY AT THE SYMPOSIUM IN CHICAGO ON MARCH 25, 1921.

BY PROFESSOR DUNHAM JACKSON.

1. *Introduction.* A great part of the progress made in mathematical analysis during the last hundred years has been closely related, either logically or historically, to the study of Taylor's and Fourier's series. The power series is fundamental, or can be made fundamental, for the theory of functions of complex variables, and the theory of functions of real variables has been elevated to its present dignity and scope largely by the successive additions made to meet the demands of the inquirer into the properties of trigonometric series.*

One of the outlying portions of the structure is built around the problem of the approximate representation of an arbitrary function by means of polynomials or by means of finite trigonometric sums in general; that is, with the admission of coefficients other than those of a specified number of terms of a particular series. This theory, which has grown to its present extent mainly within twenty years, forms the subject of the following report. It can not be set off from the rest by any sharp line of demarcation, but there are certain well-marked processes of development which it is not difficult to trace. It has seemed expedient for various reasons to make this paper an introduction to the literature of the subject, rather than an independent exposition of any considerable part of the theory. Even with this limitation, the treatment is merely illustrative, not in any sense exhaustive.

2. *Tchebychef's Theory and its Generalizations.* Between 1850 and 1860, Tchebychef † discussed the problem of de-

* Cf. Van Vleck, *The influence of Fourier's series upon the development of mathematics*, SCIENCE, vol. 39 (1914), pp. 113–124.

† Tchebychef 1 (this form of reference will be adopted when there is occasion to cite more than one paper by the same author), *Théorie des mécanismes connus sous le nom de parallélogrammes*, MÉMOIRES PRÉSENTÉS À L'ACADÉMIE IMPÉRIALE DES SCIENCES DE ST.-PÉTERSBOURG PAR DIVERS SAVANTS, vol. 7 (1854), pp. 539–568; *Oeuvres*, vol. 1, Petrograd, 1899, pp. 111–143.

Tchebychef 2, *Sur les questions de minima qui se rattachent à la représenta-*

termining a polynomial $P_n(x)$, of given degree n, to approximate a given continuous function $f(x)$, in such a way that the maximum of the absolute value of the error,

$$\text{Max. } |f(x) - P_n(x)|,$$

shall be as small as possible. His reasoning, as recorded at that early date, was naturally incomplete, according to present standards. It was put into modern form by Kirchberger* in 1902. The idea was then rapidly carried further. It was applied by Fréchet† and by J. W. Young‡ to problems of a considerably higher degree of generality, and in particular to representation by finite trigonometric sums, by Tonelli§ to functions of more than one variable (a phase already touched upon by Kirchberger and Fréchet), and by Sibirani‖ to representation by linear combinations of a given set of linearly independent functions generally; and it has been extended in a variety of other ways.

The most striking facts with regard to the Tchebychef polynomial of given degree n, for a given continuous function $f(x)$, in a given interval $a \leqq x \leqq b$, are that it is uniquely determined, and that the error $f(x) - P_n(x)$ takes on its greatest numerical value not just once, but at least $n + 2$ times, alternately with opposite signs.¶ When $n = 0$, for

---

tion approximative des fonctions, MÉMOIRES DE L'ACADÉMIE IMPÉRIALE DES SCIENCES DE ST.-PETERSBOURG, (6), SCIENCES MATHÉMATIQUES ET PHYSIQUES, vol. 7, (1859), pp. 199–291; *Oeuvres*, vol. 1, pp. 273–378.

In preparing the present paper, I have consulted only the collected works, from which the citations of the original memoirs are quoted.

* Kirchberger, *Ueber Tchebychefsche Annäherungsmethoden*, Dissertation, Göttingen, 1902. See also Blichfeldt, *Note on the functions of the form* $f(x) \equiv \varphi(x) + \alpha_1 x^{n-1} + \alpha_2 x^{n-2} + \cdots + \alpha_n$ *which in a given interval differ the least possible from zero*, TRANSACTIONS OF THE AMERICAN MATH. SOCIETY, vol. 2 (1901), pp. 100–102.

† Fréchet 1, *Sur l'approximation des fonctions par des suites trigonométriques limitées*, COMPTES RENDUS, vol. 144 (1907), pp. 124–125; Fréchet 2, *Sur l'approximation des fonctions continues périodiques par les sommes trigonométriques limitées*, ANNALES DE L'ECOLE NORMALE SUPÉRIEURE, (3), vol. 25 (1908), pp. 43–56.

‡ J. W. Young, *General theory of approximation by functions involving a given number of arbitrary parameters*, TRANSACTIONS OF THE AMERICAN MATH. SOCIETY, vol. 8 (1907), pp. 331–344.

§ Tonelli, *I polinomi d'approssimazione di Tchebychev*, ANNALI DI MATEMATICA, (3), vol. 15 (1908), pp. 47–119.

‖ Sibirani, *Sulla rappresentazione approssimata delle funzioni*, ANNALI DI MATEMATICA, (3), vol. 16 (1909), pp. 203–221.

¶ In the paper of 1854, the earliest reference to the subject with which I am acquainted, Tchebychef says (*Oeuvres*, vol. 1, p. 114):

"Soit fx une fonction donnée, U un polynome du degré n avec des coefficients arbitraires. Si l'on choisit ces coefficients de manière à ce que la example, it is evident that the best constant is that midway between the greatest and least values of $f(x)$, so that the maximum deviation is reached at least twice, once positively and once negatively. A little experimentation with a ruler and a curve drawn on paper will leave little doubt as to the correctness of the statement for $n = 1$, and perhaps will suggest at the same time the idea of the general proof. The latter, however, is by no means trivial, and it is a very satisfactory exercise in the application of elementary theorems of algebra and of the simplest principles of analysis.

For the corresponding trigonometric problem, let it be supposed that $f(x)$ is of period 2π, and is continuous for all real values of x. Among all expressions of the form

$$T_n(x) = a_0 + a_1 \cos x + \cdots + a_n \cos nx \\ + b_1 \sin x + \cdots + b_n \sin nx,$$

that is, among all finite trigonometric sums of order n, there will be one and just one for which the maximum of the quantity $|f(x) - T_n(x)|$ has the smallest possible value. This particular sum is characterized by the fact that the quantity $f(x) - T_n(x)$ reaches its greatest numerical value at least $2n + 2$ times in a period, alternately with opposite signs. It will be seen that the number $2n + 2$ here, like the number $n + 2$ in the polynomial case, is one more than the number of arbitrary coefficients in question.

Except for certain specific references in a later section, there is perhaps no occasion to dwell upon the subject longer here, since its main features are very readably presented in standard treatises.* The explicit determination of the approximating function, or of the degree of approximation attained, is extraordinarily difficult, even in relatively simple cases, because the dependence of the approximating function on $f(x)$ is not linear. It is especially noteworthy that the Tchebychef

différence $fx - U$, depuis $x = a - h$, jusqu'à $x = a + h$, reste dans les limites les plus rapprochées de 0, la différence $fx - U$ jouira, *comme on le sait*, de cette propriété:

'Parmi les valeurs les plus grandes et les plus petites de la différence $fx - U$ entre les limites $x = a - h$, $x = a + h$, on trouve au moins $n + 2$ fois la même valeur numérique.'"

The italics are mine. Kirchberger, op. cit., p. 6, states that the problem was originally proposed by Poncelet.

* Cf., e.g., Borel, *Leçons sur les Fonctions de Variables Réelles et les Développements en Séries de Polynomes*, Paris, 1905, pp. 82–92; de la Vallée Poussin 1, *Leçons sur l'Approximation des Fonctions d'une Variable Réelle*, Paris, 1919, see chapters 6, 7.

theory did not lead in any direct way to a recognition of the fundamental fact brought out in the next section, that there exist uniformly convergent processes of approximation by polynomials or by finite trigonometric sums, for an arbitrary continuous function $f(x)$, although the approximating functions of Tchebychef by definition yield the most rapidly convergent of all such processes.

3. *Weierstrass's Theorem.* Let $f(x)$ be an arbitrary continuous function over the interval $a \leqq x \leqq b$, and let ϵ be an arbitrary positive quantity. Then there will always exist a polynomial $P(x)$ such that $|f(x) - P(x)| < \epsilon$ throughout the interval. This theorem was first published by Weierstrass* in 1885. Almost at the same time, and independently of Weierstrass, Runge,† without formulating this particular conclusion, supplied the materials for a second proof, quite different in form. Other writers successively added other demonstrations, in great variety.‡ A time came when there was no longer any distinction in inventing a proof of Weierstrass's theorem, unless the new method could be shown to possess some specific excellence, in the way of simplicity, for example, or in rapidity of convergence. The question of degree of convergence will be the principal concern of the remainder of the paper. In respect to simplicity, mention may be made particularly of the proof of Lebesgue,§ which depends on the application of the binomial theorem to the representation of a function whose graph is a broken line, the proof of Landau,‖ which makes use of an especially con-

* Weierstrass 1, *Über die analytische Darstellbarkeit sogenannter willkürlicher Functionen einer reellen Veränderlichen,* BERLINER SITZUNGSBERICHTE, 1885, pp. 633–639.

† Runge, *Über die Darstellung willkürlicher Functionen,* ACTA MATHEMATICA, vol. 7 (1885), pp. 387–392, together with an earlier paper by the same author: *Zur Theorie der eindeutigen analytischen Functionen,* ACTA MATHEMATICA, vol. 6 (1885), pp. 229–244; pp. 236–237. The first-named paper deals with the approximation of an arbitrary continuous function by means of a rational function, while the other supplies the necessary facts about the approximate representation of rational functions by means of polynomials.

‡ Cf., e.g., Borel, op. cit., pp. 50–61.

§ Lebesgue 1, *Sur l'approximation des fonctions,* BULLETIN DES SCIENCES MATHÉMATIQUES, (2), vol. 22 (1898), pp. 278–287; cf. de la Vallée Poussin 1, pp. 3–5.

‖ Landau, *Über die Approximation einer stetigen Funktion durch eine ganze rationale Funktion,* RENDICONTI DEL CIRCOLO MATEMATICO DI PALERMO, vol. 25 (1908), pp. 337–345.

venient form of (so-called *singular*) definite integral, and Simon's modification* of Landau's proof, in which the definite integral is replaced by a finite sum.

The trigonometric form of the theorem, to the effect that an arbitrary continuous function of period 2π can be uniformly represented by a finite trigonometric sum with any assigned degree of accuracy, was established by Weierstrass himself, not in the first paper already cited, but immediately afterwards.† The polynomial and trigonometric cases are not only susceptible of parallel treatment, but are readily convertible one into the other by a simple change of variable. Among the direct proofs for the trigonometric case may be mentioned, for simplicity and elegance, those of de la Vallée Poussin,‡ using a definite integral, and Kryloff,§ using the finite sum which corresponds to de la Vallée Poussin's integral.

4. *First Studies of Degree of Convergence. De la Vallée Poussin's Problem.* It has long been known, in more or less detail, that there is a relation between the properties of continuity of a function and the degree of accuracy that can be attained in its approximate representation by specified means. For example, Picard,|| in his *Traité d'Analyse,* points out incidentally that if $f(x)$ is a function of period 2π possessing a kth derivative which is (essentially) of limited variation, the coefficient of cos nx or sin nx in the Fourier series for $f(x)$ does not exceed a constant multiple of $1/n^{k+1}$ in absolute value. This leads to a theorem about the degree of approximation to $f(x)$ given by the partial sum of its Fourier series, that is by a particular trigonometric sum of the nth order.

Lebesgue,¶ in 1908, formally proposed the problem of discussing the relation between the accuracy of polynomial ap-

---

*Simon, *A formula of polynomial interpolation*, ANNALS OF MATHEMATICS, (2), vol. 19 (1918), pp. 242–245.

† Weierstrass 2, same title as his first paper, BERLINER SITZUNGSBERICHTE, 1885, pp. 789–805.

‡ de la Vallée Poussin 2, *Sur l'approximation des fonctions d'une variable réelle et de leurs dérivées par des polynômes et des suites limitées de Fourier*, BULLETINS DE L'ACADÉMIE ROYALE DE BELGIQUE, Classe des Sciences, 1908, pp. 193–254.

§ Kryloff, *Sur quelques formules d'interpolation généralisée*, BULLETIN DES SCIENCES MATHÉMATIQUES, (2), vol. 41 (1917), pp. 309–320.

|| Picard, *Traité d'Analyse*, 2nd ed., vol. 1, pp. 252–253, 255–256.

¶ Lebesgue 2, *Sur la représentation approchée des fonctions*, RENDICONTI DEL CIRCOLO MATEMATICO DI PALERMO, vol. 26 (1908), pp. 325–328.

proximation and the requisite degree of the polynomial, and indicated that if $f(x)$ satisfies a Lipschitz condition for $a \leqq x \leqq b$, it can be approximately represented by a polynomial of the nth degree with an error not exceeding a constant multiple of $\sqrt{(\log n)/n}$. This result was improved by de la Vallée Poussin,* who obtained the limit $1/\sqrt{n}$ in place of that just mentioned.† For a more restricted class of functions, retaining enough generality to include any function whose graph is a broken line with a finite number of segments, he found‡ the still closer limit $1/n$. Then he added,§

"Il serait très intéressant de savoir s'il est impossible de représenter l'ordonnée d'une ligne polygonale avec une approximation d'ordre supérieur à $1:n$ par un polynôme de degré n."

This remark has been the direct or indirect occasion of most of the subsequent work on the subject.

5. *Inner Limit of Approximation. S. Bernstein's Theory.* The simplest example of a function coming within the specifications of the problem is the function $|x|$, considered in the interval from -1 to 1. It was shown by de la Vallée Poussin‖ that the maximum error of an approximating polynomial of degree n can not approach zero faster than $1/(n \log^3 n)$. S. Bernstein¶ and the writer** independently replaced this limit by $1/(n \log n)$. The final solution, verifying de la Vallée Poussin's surmise that $1/n$ is the actual limit, was given by S. Bernstein,†† in a notable prize essay for the Belgian Academy,

---

* de la Vallée Poussin 2, already cited, p. 222.

† The hypothesis was somewhat differently stated by de la Vallée Poussin, but his analysis is immediately applicable to the case of a Lipschitz condition; cf. D. Jackson 1, cited below, p. 10, footnote.

‡ de la Vallée Poussin 3, *Note sur l'approximation par un polynôme d'une fonction dont la dérivée est à variation bornée*, BULLETINS DE L'ACADÉMIE ROYALE DE BELGIQUE, Classe des Sciences, 1908, pp. 403–410.

§Footnote, p. 403.

‖ de la Vallée Poussin 4, *Sur les polynomes d'approximation et la représentation approchée d'un angle*, BULLETINS DE L'ACADÉMIE ROYALE DE BELGIQUE, Classe des Sciences, 1910, pp. 808–844.

¶ S. Bernstein 1, *Sur l'approximation des fonctions continues par des polynomes*, COMPTES RENDUS, vol. 152 (1911), pp. 502–504.

** D. Jackson 1, *Über die Genauigkeit der Annäherung stetiger Funktionen durch ganze rationale Funktionen gegebenen Grades und trigonometrische Summen gegebener Ordnung*, Dissertation, Göttingen, 1911; see pp. 49–52.

†† S. Bernstein 2, *Sur l'ordre de la meilleure approximation des fonctions continues par des polynomes de degré donné*, Mémoire couronné, Brussels, 1912. (Included in MÉMOIRES PUBLIÉS PAR LA CLASSE DES SCIENCES DE

with de la Vallée Poussin's problem for its text. Expanding into a discussion of a wide range of questions in the theory of polynomial approximation generally, the essay contains a number of results comparable in interest with the accomplishment of its original purpose. One of these, by reason of its simplicity and its far-reaching consequences, may well be regarded as one of the most remarkable theorems of recent times. It will be worth while to dwell on it at some length. Here, again, it is possible to speak either of polynomials or of trigonometric sums. The statement is more striking in terms of the latter. For the trigonometric case, the theorem is as follows:

Let $T_n(x)$ be an arbitrary trigonometric sum of order n, and let L be the maximum of its absolute value. Then the maximum of the absolute value of the derivative $T_n'(x)$ can not exceed nL.

For this formulation, to be sure, the credit is not undivided. Bernstein's own statement* asserts merely that $|T_n'(x)|$ can not attain the value $2nL$; his proof is far from simple; and its validity has been called in question,† though this last remark applies only to his discussion of the trigonometric theorem, not to the polynomial case, in which he was primarily interested. The theorem was approached by subsequent writers from different angles, and was finally revealed in its true simplicity by de la Vallée Poussin,‡ whose demonstration is well worth repeating here.

An equivalent form of the assertion to be proved is as follows:

If the maximum of $|T_n'(x)|$ is 1, the maximum of $|T_n(x)|$ can not be less than $1/n$.

Suppose the maximum of $|T_n(x)|$ were less than $1/n$. Then, for any value of the constant C, the function

$$R_n(x) = \frac{1}{n}\sin(nx + C) - T_n(x)$$

would have the sign of $\sin(nx + C)$ at each of the $2n$ points

L'ACADÉMIE ROYALE DE BELGIQUE, (2), vol. 4; the detailed citations below, however, refer to the page-numbers of the essay itself, as printed separately.)

* S. Bernstein 2, p. 20.

† Cf. de la Vallée Poussin, passages cited in next footnote.

‡ de la Vallée Poussin 5, *Sur le maximum du module de la dérivée d'une expression trigonométrique d'ordre et de module bornés*, COMPTES RENDUS, vol. 166 (1918), pp. 843–846; de la Vallée Poussin 1, pp. 39–42.

at which $|\sin(nx + C)|$ has a maximum, in an interval of length 2π. Because of the alternation of these signs, and the periodicity of the functions involved, $R_n(x)$ would then vanish for at least $2n$ distinct values of x in a period, and consequently, by Rolle's theorem, the same thing would be true of the derivative

$$R_n'(x) = \cos(nx + C) - T_n'(x).$$

But if C is chosen so as to make $\cos(nx + C)$ coincide in value with $T_n'(x)$ at a point where the latter is equal to ± 1, $R_n'(x)$ will have a double root at this point. Hence it will have, in an entire period, roots of aggregate multiplicity at least $2n + 1$. For a trigonometric sum of order n, which can not vanish identically—since $R_n(x)$, by reason of its changes of sign, can not be a constant—this is impossible, and the contradiction proves the theorem. It can be shown further,* though not quite so simply, that the maximum of $|T_n(x)|$ will be *greater* than $1/n$, unless $T_n(x)$ has precisely the form $(1/n)\sin(nx + C)$, that is, the latter is the only function for which the limit specified in the theorem is attained.

The corresponding theorem for polynomials may be stated as follows, for a particular interval:

Let $P_n(x)$ be an arbitrary polynomial of degree n, and let L be the maximum of its absolute value for $-1 \leqq x \leqq 1$. Then the maximum of $|\sqrt{1 - x^2}\, P_n'(x)|$ can not exceed nL in the interval specified.

Bernstein established this fact first,† and went from it to the trigonometric case. The relative simplicity of the opposite procedure, on the basis of de la Vallée Poussin's trigonometric proof, has been pointed out by the writer.‡

The most immediate application of the theorem is to a type of argument of which the following is a simple case. Let $f(x)$ be a continuous function of period 2π, and let it be supposed that there exists, for every positive integral value of n, a trigonometric sum $T_n(x)$, of order n, so that

$$|f(x) - T_n(x)| \leqq Q/n^2,$$

where Q is independent of n and x. Then $f(x)$ is the sum of

* Cf. citations in preceding footnote.

† S. Bernstein 2, pp. 6–11.

‡ D. Jackson 2, *On the convergence of certain trigonometric and polynomial approximations*, TRANSACTIONS OF THE AMERICAN MATH. SOCIETY, vol. 22 (1921), pp. 158–166; see p. 162.

the uniformly convergent series

$$T_1 + (T_2 - T_1) + (T_4 - T_2) + (T_8 - T_4) + \cdots,$$

or, if $f(x) - T_n(x) = r_n(x)$,

(1) $f(x) = T_1 + (r_1 - r_2) + (r_2 - r_4) + (r_4 - r_8) + \cdots.$

It will be shown that if the series on the right is differentiated term by term, the resulting series will be uniformly convergent, and $f(x)$ consequently must have a continuous first derivative. Since $|r_n| \leqq Q/n^2$, by hypothesis,

$$|r_1 - r_2| \leqq Q + \frac{Q}{2^2} < 2Q, \quad |r_2 - r_4| \leqq \frac{Q}{2^2} + \frac{Q}{4^2} < \frac{2Q}{2^2},$$
$$|r_4 - r_8| < \frac{2Q}{4^2}, \quad \cdots.$$

But since $r_1 - r_2$ is a trigonometric sum of order 2, etc., it follows from the theorem on the derivative of a trigonometric sum that

$$|(r_1 - r_2)'| < 2 \cdot 2Q, \quad |(r_2 - r_4)'| < 4 \cdot \frac{2Q}{2^2}, \quad \cdots,$$

and, generally, that the derivative of the $(p+1)$th term on the right in (1) does not exceed

$$2^p \cdot \frac{2Q}{2^{2(p-1)}} = \frac{Q}{2^{p-3}}.$$

The last quantity is the general term of a convergent series, and the uniform convergence of the series of derivatives is established. It is evident that this illustration by no means exhausts the possibilities of the method. Thus we may state the following more general theorem.*

If $f(x)$ can be represented by trigonometric sums of order n with an error not exceeding $Q/n^{k+\alpha}$, where k is a positive integer or zero, and $0 < \alpha < 1$, then $f(x)$ has a continuous kth derivative satisfying a Lipschitz condition of order α, that is

$$|f^{(k)}(x') - f^{(k)}(x'')| \leqq \lambda |x' - x''|^\alpha,$$

where λ is a constant.† The idea can be carried still further.

* Cf. de la Vallée Poussin 1, p. 57; also S. Bernstein 1, and S. Bernstein 2, pp. 22–23, 27.

† The value $\alpha = 1$ is ruled out in the statement of the theorem; but if the hypothesis is satisfied for $\alpha = 1$, as in the preceding illustration ($k = 1$), it is of course satisfied, and the conclusion holds, for an arbitrary value of $\alpha < 1$.

The theorem concerning the order of approximation to $|x|$, like the theorem on the derivative of a trigonometric sum, is most accessible at present through the exposition of de la Vallée Poussin.* The idea of his proof is as follows. In the first place, the problem is referred to the equivalent one of representing $|\sin x|$ by a trigonometric sum of order n in x. Lebesgue† had pointed out that no trigonometric sum of order n can give an error smaller than the corresponding remainder in the Fourier series, multiplied by a quantity of the order of $1/\log n$. As it is readily recognized that the error in the Fourier series for $|\sin x|$ is of the order of $1/n$, the quantity $1/(n \log n)$, already mentioned as a step toward the final result, is seen at once to be an inner limit for the order of the best approximation. If the Fejér mean could be used in the same way as the Fourier sum proper, the desired limit $1/n$ would be obtained with equal ease; for Lebesgue's remark is a consequence of the fact that the partial sum of the Fourier series for an arbitrary function can not exceed the maximum of the absolute value of the function itself, multiplied by a quantity of the order of $\log n$, and in the case of the corresponding Fejér mean, this multiplier is replaced by 1. But the method fails at first, because it is essential also that the Fourier sum of order n for a function $T_n(x)$ is identical with $T_n(x)$, while this is not true of the Fejér mean. However, de la Vallée Poussin observes that if $\tau_k(x)$ is the Fejér mean of order $k - 1$, $T_n(x)$ is reproduced identically by the formula

$$T_n(x) = 2\tau_{2n}(x) - \tau_n(x);$$

and from this relation he is able to draw the desired inference, not quite immediately, but by easy steps, in the course of two or three pages. The general theorem which intervenes is as follows.

Let $f(x)$ be a given arbitrary function, let $S_k(x)$ be the partial sum of its Fourier series to terms of order k, and let $T_n(x)$ be any trigonometric sum of order n. Then the maximum of $|f(x) - T_n(x)|$ can not be less than one-fourth the maximum of

$$\left| f(x) - \frac{S_n + S_{n+1} + \cdots + S_{2n-1}}{n} \right|.$$

* de la Vallée Poussin 6, *Sur la meilleure approximation des fonctions d'une variable réelle par des expressions d'ordre donné*, Comptes Rendus, vol. 166 (1918), pp. 799–802; de la Vallée Poussin 1, pp. 33–37.

† Cf. e.g. Lebesgue 3, *Sur les intégrales singulières*, Annales de la Faculté de Toulouse, (3), vol. 1 (1909), pp. 25–117; see pp. 116–117.

Bernstein himself, in the prize essay, proves the theorem about the polynomial representation of $|x|$ in two different ways.* Both of his demonstrations depend on an extension, interesting in itself, of the Tchebychef theory to the problem of the approximate representation of a given continuous function by linear combinations of powers of x with arbitrarily given exponents.† The exponents may be any positive real numbers, not necessarily integral. There is a real generalization, however, even if they are merely selected integers; there would be no point in saying that some powers *may* be omitted, in the approximating polynomials used, for that is always understood, as a matter of course; but the problem is essentially changed if it is demanded that certain powers *shall* be omitted.

By way of obtaining a corresponding generalization of Weierstrass's theorem, Bernstein further inquires under what circumstances a sequence of positive powers x^{α_1}, x^{α_2}, $\cdots$, will be sufficient (in technical language, *complete*) for the uniformly convergent representation of an arbitrary continuous function.‡ He derives a condition which is necessary, and others which are sufficient. Müntz and Szász§ later established a condition which is both necessary and sufficient, namely (except for minor qualifications) the very simple requirement that the series $\Sigma\,(1/\alpha_n)$ diverge. The last-named author extends the discussion to the case of complex exponents.

The rest of Bernstein's memoir must be dismissed with the utmost brevity in the present summary. Among the topics treated may be mentioned some results concerning the absolute value of a polynomial for complex values of the argument,|| a study of the approximate representation of analytic functions of a complex variable,¶ theorems with regard to the

* S. Bernstein 2, pp. 55–62. See also S. Bernstein 3, *Sur la meilleure approximation de $|x|$ par des polynomes de degrés donnés*, ACTA MATHEMATICA, vol. 37 (1914), pp. 1–57.

† S. Bernstein 2, pp. 38–41.

‡ S. Bernstein 2, pp. 78–84.

§ Szász, *Über die Approximation stetiger Funktionen durch lineare Aggregate von Potenzen*, MATHEMATISCHE ANNALEN, vol. 77 (1916), pp. 482–496. The paper of Müntz, which I have not seen, is cited by Szász, loc. cit., p. 483, footnote.

|| pp. 13–15.

¶ pp. 36, 65–76, 94–95, and elsewhere. This subject, which has an extensive literature of its own, must be left outside the scope of the present review.

degree of convergence of Fourier's series and the Fejér means of Fourier's series,* in the spirit of the following section, and applications of the theory of approximation to questions of the existence of derivatives of functions of more than one variable.†

A paper by Montel‡ presents a noteworthy continuation of Bernstein's work in various directions, particularly with reference to derivatives of fractional order.

6. *Outer Limit of Approximation.* From Picard's inequalities, already cited, for the coefficients in a Fourier series, it follows that, if the function developed has a continuous kth derivative which is of limited variation, the remainder after n terms of the series does not exceed a constant multiple of $1/n^k$. For $k = 1$, the same outer limit of error for the approximation obtainable by means of finite trigonometric sums of order n was found by de la Vallée Poussin,§ under a similar but somewhat more general hypothesis. Most often, however, such discussion has involved the hypothesis of a Lipschitz condition, either on the function itself or on one of its derivatives.

For the case of a function satisfying a Lipschitz condition, de la Vallée Poussin,‖ as has already been noted, found the order of approximation $1/\sqrt{n}$, and the same limit was recorded a little later by Lebesgue,¶ both for polynomials and for trigonometric sums. From the point of view of a rapid survey of the problem, it is not necessary to specify each time which kind of approximating function is used, as the two classes of results are to a large extent interchangeable.

Another step in advance was presently taken by Lebesgue,** who proved that the remainder in the Fourier series for a function of the kind under discussion can not exceed a constant multiple of $(\log n)/n$.

* pp. 85–93.

† pp. 97–103.

‡ Montel, *Sur les polynomes d'approximation*, BULLETIN DE LA SOCIÉTÉ MATHÉMATIQUE DE FRANCE, vol. 46 (1919), pp. 151–192.

§ de la Vallée Poussin 3.

‖ de la Vallée Poussin 2, p. 222.

¶ Lebesgue 3, pp. 112–115.

** Lebesgue 4, *Sur la représentation trigonométrique approchée des fonctions satisfaisant à une condition de Lipschitz*, BULLETIN DE LA SOCIÉTÉ MATHÉMATIQUE DE FRANCE, vol. 38 (1910), pp. 184–210; pp. 199–201.

He also showed that this result is final, as far as the Fourier series is concerned, that is that there exist functions satisfying a Lipschitz condition, for which the remainder does not approach zero any more rapidly.*

There still remained a question as to the degree of approximation that might be attained by a different choice of the approximating functions. The writer showed that a function satisfying a Lipschitz condition can be represented by a trigonometric sum of the nth order,† or by a polynomial of the nth degree,‡ with a maximum error not exceeding a constant multiple of $1/n$, and that the limit $1/n$ in this general statement can not be replaced by any infinitesimal of higher order.§ The truth of the last part of this assertion of course also follows immediately from Bernstein's results with regard to the function $|x|$, which were published a little later. The writer subsequently obtained inequalities for the magnitude of the numerical constants involved, so that the following statements can be made in summary.||

If $f(x)$ is a function of period 2π satisfying everywhere the condition

$$|f(x') - f(x'')| \leqq \lambda |x' - x''|, \tag{2}$$

it is possible to represent $f(x)$ by a trigonometric sum of order n with a maximum error not exceeding $K\lambda/n$, where K is an absolute constant; the statement is true for $K = 3$, but not for any value of K less than $\pi/2$.

If $f(x)$ satisfies the condition (2) for $a \leqq x \leqq b$, it can be represented throughout this interval by a polynomial of the nth degree with a maximum error not exceeding $L\lambda(b - a)/n$, where L is an absolute constant; the statement is true for $L = 1\frac{1}{2}$, but not for any value of L less than $\frac{1}{2}$.

The upper values for K and L, as actually computed by the writer, were slightly less than 3 and $1\frac{1}{2}$ respectively, and they can be further reduced by other methods, due particularly to

* Lebesgue 4, pp. 202–206.

† D. Jackson 1, Theorem VI; see also top of p. 46.

‡ D. Jackson 1, Theorems I, IV.

§ D. Jackson 1, Theorems XII, XIII.

|| D. Jackson 3, *On approximation by trigonometric sums and polynomials*, TRANSACTIONS OF THE AMERICAN MATH. SOCIETY, vol. 13 (1912), pp. 491–515; Theorems I, II, VI, VII. There is a misprint in the statement of Theorem VII; for L_2 should be read L_1.

Gronwall.* The ultimate determination of the best values for K and L, however, is still an open question.

The order of approximation being once established for the case of a Lipschitz condition, corresponding results can be deduced without difficulty both for more general and for more restrictive hypotheses. By way of generalization,† let $f(x)$ be an arbitrary continuous function, of period 2π, say, to restrict attention to the trigonometric case, and let $\omega(\delta)$ be the maximum of $|f(x')-f(x'')|$ for $|x'-x''| \leqq \delta$. Let $f_1(x)$ be a function represented graphically by a broken line, equal to $f(x)$ at the points $x = 2j\pi/n$ $(j = 0, \pm 1, \pm 2, \cdots)$, and varying linearly from one of these points to the next. Then, for any x, $f(x)$ and $f_1(x)$ will each differ from the value at the nearest vertex point by not more than $\omega(2\pi/n)$, and

$$|f(x) - f_1(x)| \leqq 2\omega(2\pi/n)$$

for all values of x. Since $f_1(x)$ satisfies a Lipschitz condition with coefficient $\omega(2\pi/n)/(2\pi/n)$, there will exist trigonometric sums $T_n(x)$, of order n, so that

$$|f_1(x) - T_n(x)| \leqq 3\cdot\frac{\omega(2\pi/n)}{2\pi/n}\cdot\frac{1}{n} = \frac{3}{2\pi}\,\omega(2\pi/n),$$

and

$$|f(x) - T_n(x)| \leqq \left(\frac{3}{2\pi}+2\right)\omega\left(\frac{2\pi}{n}\right).$$

A similar result can also be obtained directly by means of a definite integral, without the intermediate reference to the Lipschitz condition.‡

On the other hand, suppose that $f(x)$, once more assumed to

---

* Unpublished letters to the present writer, January 21 and February 16, 1913; *On approximation by trigonometric sums*, this BULLETIN, vol. 21 (1914–15), pp. 9–14. See also de la Vallée Poussin 1, pp. 44–46. My computation, in the Transactions article cited, involved the ratio of two integrals, J_m' and J_m; within a few weeks after its publication, I was in receipt of letters from Messrs. Gronwall, I. Schur, and D. Cauer (the last-named, then a student at Göttingen, writing informally in behalf of Professor Landau), which contained, among other data, two proofs that

$$J_m = \frac{\pi}{3}\left(1+\frac{1}{2m^2}\right),$$

and seven proofs (including one quoted by Schur from Marcel Riesz) that $\lim_{m\to\infty} J_m' = \log 2$.

† See, e.g., D. Jackson 1, Theorem VIII. The fundamental idea of the proof is derived from Lebesgue 4, p. 202.

‡ Cf. de la Vallée Poussin 1, pp. 44–46.

have the period 2π, is not merely continuous itself, but has a continuous first derivative satisfying the Lipschitz condition

$$|f'(x') - f'(x'')| \leqq \lambda |x' - x''|.$$

There is a finite trigonometric sum $T_n'(x)$ such that

$$|f'(x) - T_n'(x)| \leqq \frac{3\lambda}{n}.$$

Moreover, it follows from the proof of the theorem on which this assertion is based, though not from the statement of the theorem itself, that T_n' can be taken so that its constant term is zero,* and its integral therefore is also a trigonometric sum of order n. Let the expression

$$f(0) + \int_0^x T_n'(x)dx$$

be denoted by $T_n(x)$, and let $f'(x) - T_n'(x) = r_n(x)$. Then

$$f(x) = T_n(x) + \int_0^x r_n(x)dx.$$

Of course the integral on the right does not exceed a constant multiple of $1/n$, but it is possible to say much more than that. For the integral has a derivative which does not exceed $3\lambda/n$, that is the integral itself *satisfies a Lipschitz condition with coefficient* $3\lambda/n$, *and can be represented by a trigonometric sum of order n with an error not exceeding* $9\lambda/n^2$. This sum, combined with $T_n(x)$, gives an equally close approximation for $f(x)$. In general†—though the proof requires a little attention in detail—*if* $f(x)$ *has a* $(k-1)$*th derivative which satisfies a Lipschitz condition with coefficient* λ, *then* $f(x)$ *can be approximately represented by a trigonometric sum of order n with an error not exceeding* $3^k\lambda/n^k$. There is a corresponding but slightly less simple result for polynomial approximation.‡

From the trigonometric theorem it follows immediately, by a remark of Lebesgue already cited, that if $f(x)$ has a $(k-1)$th

* That is, if the function for which an approximate representation is sought, $f'(x)$ in this case, is such that its integral is periodic, the particular approximating function defined in the course of the demonstration will also have a periodic integral.

† D. Jackson 4, *On the approximate representation of an indefinite integral and the degree of convergence of related Fourier's series*, TRANSACTIONS OF THE AMERICAN MATH. SOCIETY, vol. 14 (1913), pp. 343–364; Theorem III. Cf. also D. Jackson 1, Theorem VII, and D. Jackson 3, Theorems III, VIII.

‡ D. Jackson 4, Theorem VII; cf. D. Jackson 1, Theorems II, IVa, and D. Jackson 3, Theorems IV, IX.

derivative satisfying a Lipschitz condition, the remainder after terms of the nth order in the Fourier series for $f(x)$ does not exceed $(\log n)/n^k$, multiplied by a quantity independent of n and x. It is noteworthy that a different method yields a still closer result here, to the extent that the constant multiplier obtained is *independent of* k. If the coefficient in the Lipschitz condition is λ, the error does not exceed $\lambda(\log n)/n^k$, multiplied by an absolute constant.*

Finally, we may note a case in which one of the theorems of the preceding section has an exact converse.† *If $f(x)$ has a continuous kth derivative satisfying a Lipschitz condition of order α, where k is a positive integer or zero, and $0 < \alpha < 1$, then $f(x)$ can be represented by trigonometric sums of order n with an error not exceeding $Q/n^{k+\alpha}$, where Q is independent of n and x.* Taken without regard to its converse, the statement can be made both more precise, by exhibiting the dependence of Q on the coefficient in the Lipschitz condition, and more general, by varying the hypothesis on the kth derivative.

7. *Trigonometric Interpolation.* In the theory which forms the subject of § 6, much use is made of formulas involving definite integrals. It is possible to vary the treatment by replacing the integrals by finite sums. From this substitution, which works out most satisfactorily in the trigonometric case, results an extensive theory of trigonometric interpolation.‡ The ordinary formula of interpolation with equidistant ordinates has of course been known for a long time, but its

---

* Cf. S. Bernstein 2, pp. 92–93; D. Jackson 4, Theorem X; de la Vallée Poussin 1, pp. 23–25, 27–29.

† Cf. de la Vallée Poussin 1, pp. 51–52, 57; also D. Jackson 4, Theorem IV.

‡ Cf., e. g., de la Vallée Poussin 7, *Sur la convergence des formules d'interpolation entre ordonnées équidistantes*, BULLETINS DE L'ACADÉMIE ROYALE DE BELGIQUE, Classe des Sciences, 1908, pp. 319–403; Faber 1, *Über stetige Funktionen*, MATHEMATISCHE ANNALEN, vol. 69 (1910), pp. 372–443, see pp. 417–443; Faber 2, *Über die interpolatorische Darstellung stetiger Funktionen*, JAHRESBERICHT DER DEUTSCHEN MATHEMATIKER-VEREINIGUNG, vol. 23 (1914), pp. 192–210; D. Jackson 5, *On the accuracy of trigonometric interpolation*, TRANSACTIONS OF THE AMERICAN MATH. SOCIETY, vol. 14 (1913), pp. 453–461; D. Jackson 6, *A formula of trigonometric interpolation*, RENDICONTI DEL CIRCOLO MATEMATICO DI PALERMO, vol. 37 (1914), pp. 371–375; D. Jackson 7, *Note on trigonometric interpolation*, RENDICONTI DEL CIRCOLO MATEMATICO DI PALERMO, vol. 39 (1915), pp. 230–232; D. Jackson 8, *On the order of magnitude of the coefficients in trigonometric interpolation*, TRANSACTIONS OF THE AMERICAN MATH. SOCIETY, vol. 21 (1920), pp. 321–332.

convergence seems to have been first studied in detail by de la Vallée Poussin and Faber within fifteen years. The convergence properties of the interpolating formula are similar to, but not absolutely identical with, those of the Fourier series, which it formally resembles. The analogy extends to questions of degree of convergence, and to the various methods of trigonometric approximation which have been devised for one special purpose or another.

8. *The Method of Least mth Powers.* Of modest importance in itself, perhaps, but of some interest as affording a field for the application of the preceding results, is the theory of what may be called the method of least mth powers.* If $f(x)$ is a continuous function of period 2π (to restrict attention to the trigonometric case once more), it is well known that the trigonometric sum $T_n(x)$, of order n, for which the integral

$$\int_0^{2\pi} [f(x) - T_n(x)]^2 dx$$

has the least possible value, is the partial sum of the Fourier series for $f(x)$. The condition determining $T_n(x)$ can be generalized by writing, in place of the square of the error, the mth power of its absolute value, where m is an arbitrary positive exponent, most conveniently assumed to be greater than 1. The resulting sum $T_n(x)$ is not readily amenable to calculation, because, like the Tchebychef sum, it does not depend linearly on $f(x)$; but it is possible to show that it always exists, is uniquely determined, for $m > 1$, and approaches the Tchebychef sum as a limit when m becomes infinite. Furthermore, if m is held fast and n is allowed to increase indefinitely, $T_n(x)$ will converge uniformly to the value $f(x)$, under suitable hypotheses. The proof depends on Bernstein's theorem concerning the derivative of a trigonometric sum, and on the possibility of representing $f(x)$ by a properly chosen sum with a specified degree of approximation.

THE UNIVERSITY OF MINNESOTA.

* Cf. D. Jackson 9, *On functions of closest approximation*, TRANSACTIONS OF THE AMERICAN MATH. SOCIETY, vol. 22 (1921), pp. 117–128, and D. Jackson 2, already cited. The polynomial case had previously been treated by Pólya, in a note with which I was not acquainted at the time of writing my own papers: *Sur un algorithme, etc.*, COMPTES RENDUS, vol. 157 (1913), pp. 840–843.