

ON APPROXIMATION BY FUNCTIONS OF GIVEN CONTINUITY*

BY

DUNHAM JACKSON

1. **Introduction.** Let $f(x)$ be a given function, continuous for $a \leq x \leq b$. Let λ be a given constant ≥ 0 , and let \mathcal{O}_λ , or, when no confusion will arise, simply \mathcal{O} , denote the family of all functions $\varphi(x)$ satisfying the condition

$$|\varphi(x_2) - \varphi(x_1)| \leq \lambda |x_2 - x_1|$$

whenever $a \leq x_1 \leq b$, $a \leq x_2 \leq b$. This note is concerned with the question of approximating $f(x)$ by means of a function of the family \mathcal{O} . The discussion is deliberately restricted in scope, for the sake of illustrating the leading ideas with the greatest practicable simplicity. It could be generalized, to a greater or less extent, in any one of a number of different directions.

2. **Existence and uniqueness of the approximating function according to the criterion of least m th powers, $m \geq 1$.** Let $\varphi(x)$ be an arbitrary function of the family \mathcal{O} . Such a function will be called for brevity "a function φ ". Let m be a given constant > 1 . Attention will be directed first to the minimizing of the integral

$$\delta = \int_a^b |f(x) - \varphi(x)|^m dx.$$

As φ ranges over the family \mathcal{O} , the value of δ has a lower limit $\gamma \geq 0$. It will be shown that there is one and just one function φ for which δ attains the value γ . It is understood that the function $f(x)$ and the constants λ and m are held fast.

If $f(x)$ itself belongs to \mathcal{O} , the value of γ is evidently zero, attained by taking $\varphi = f$, and by no other choice of φ . If f is not a function φ , let $\epsilon_1, \epsilon_2, \dots, \epsilon_k, \dots$ be a set of positive numbers approaching zero, and for each value of k let φ_k be a function φ for which

$$(1) \quad \delta < \gamma + \epsilon_k.$$

* Presented to the Society, April 14, 1923.

It may be assumed without loss of generality that

$$(2) \quad \mu \leq \varphi_k(x) \leq M$$

throughout (a, b) , where μ and M are the minimum and the maximum of $f'(x)$ in the interval. For if this is not the case originally, a function φ'_k , belonging to Φ , and satisfying the condition (1), can be defined by taking $\varphi'_k = \varphi_k$ when $\mu \leq \varphi_k \leq M$, $\varphi'_k = \mu$ when $\varphi_k \leq \mu$, and $\varphi'_k = M$ when $\varphi_k \geq M$. The functions φ_k will be considered chosen so that (2) is fulfilled. Then, being uniformly bounded (because of (2)) and equally continuous (by the definition of Φ), they form a sequence to which the theorem of Ascoli and Arzelà may be applied. There will be a sub-sequence among them which approaches uniformly a limit $\tau(x)$. It is recognized immediately that the uniform limit of a sequence of functions φ is a function φ . On the other hand, the values of δ corresponding to a uniformly convergent sequence of functions φ approach the value of δ corresponding to the limiting function of the sequence. As the limit of δ for the sequence (φ_k) , or any sub-sequence chosen from it, is γ , because of (1), it appears that $\tau(x)$ is a function of the family Φ , for which*

$$\int_a^b |f - \tau|^m dx = \gamma.$$

Such a function will be called briefly an *approximating function* for $f(x)$, corresponding to the parameters λ and m .

Suppose τ_1 and τ_2 are two functions φ , for each of which $\delta = \gamma$. Let $\tau_3 = \frac{1}{2}(\tau_1 + \tau_2)$. Then τ_3 is also a function φ . Furthermore, $f - \tau_3 = \frac{1}{2}[(f - \tau_1) + (f - \tau_2)]$, and

$$|f - \tau_3|^m \leq \frac{1}{2}[|f - \tau_1|^m + |f - \tau_2|^m],$$

the relation being an actual inequality at any point where $\tau_1 \neq \tau_2$. So if τ_1 and τ_2 are not identical, the value of δ corresponding to τ_3 is less than γ , which is impossible. This proves the uniqueness of the approximating function τ .

The approximating function for given λ and m will be denoted henceforth by $\varphi_{\lambda m}(x)$, instead of $\tau(x)$.

* Somewhat more directly, the existence of a minimum follows almost immediately from a well known theorem in the theory of functions of lines.

The above discussion has assumed that $m > 1$. The proof of the existence of at least one approximating function applies without change, if $0 < m \leq 1$. Examples show* that when $m < 1$ there may be more than one approximating function for a given value of m . In the limiting case $m = 1$, the approximating function is unique once more, though a modified form of demonstration is required.

Suppose, if possible, that τ_1 and τ_2 are two distinct approximating functions corresponding to the exponent $m = 1$, and let $\tau_3 = \frac{1}{2}(\tau_1 + \tau_2)$. It is still true that

$$|f - \tau_3| \leq \frac{1}{2} [|f - \tau_1| + |f - \tau_2|],$$

but the equality sign is not necessarily ruled out, even at points where $\tau_1 \neq \tau_2$. The relation would be an inequality, however, and a contradiction would result, if $f - \tau_1$ and $f - \tau_2$ were to take on *opposite signs* at any point. It must be assumed that $(f - \tau_1)(f - \tau_2) \geq 0$ throughout (a, b) . This being the case, let a function τ_4 be defined at each point of (a, b) as equal to τ_1 or τ_2 , according as $|f - \tau_1| \leq$ or $\geq |f - \tau_2|$. Then τ_4 is continuous, belongs to the family Φ , and gives a better approximation, according to the exponent $m = 1$, than either τ_1 or τ_2 , under the assumption that the latter functions are not identical. Thus a contradiction is obtained in the present circumstances.

3. Existence and indeterminacy of the approximating function according to the criterion of least maximum error. For an arbitrary φ , let d be the maximum of $|f - \varphi|$ in the interval $a \leq x \leq b$. The values of d corresponding to the various members of the family Φ have a lower limit ≥ 0 , which may be denoted by c . By reasoning altogether parallel to that of the preceding section, it may be shown that there exists a function τ in the family Φ , for which d actually has the value c .

This function τ , however, is not in general uniquely determined. For example, let $f(x) = \sqrt{x}$ in the interval $0 \leq x \leq 1$, and let $\lambda = 1$. Here the value of c can not be less than $\frac{1}{8}$. For as x goes from 0 to $\frac{1}{4}$, $f(x)$ increases from 0 to $\frac{1}{2}$, while no function φ can increase by more than $\frac{1}{4}$, so that any φ must differ from f by at least $\frac{1}{8}$ either for $x = 0$ or for $x = \frac{1}{4}$. On the other hand, d is actually equal to $\frac{1}{8}$ for the function $\tau_1(x) = x + \frac{1}{8}$, and for the function $\tau_2(x)$ which is equal to $x + \frac{1}{8}$ for $0 \leq x \leq \frac{1}{8}(3 + 2\sqrt{2})$ and to \sqrt{x}

* E. g., let $f(x)$ be defined for $-1 \leq x \leq 1$ as equal to -1 when $x \leq -\epsilon$, equal to $+1$ when $x \geq \epsilon$, and equal to x/ϵ when $-\epsilon \leq x \leq \epsilon$. Let $\lambda = 0$, $m = \frac{1}{2}$. The functions φ are constants. The approximating constant C can not be unique unless $C = 0$, for C and $-C$ give the same value to the integral δ . But $C = 0$ does not give the best approximation, when ϵ is taken arbitrarily small, because the values of δ corresponding to $C = 0$ and $C = +1$ are arbitrarily near to 2 and $\sqrt{2}$ respectively.

for $\frac{1}{8}(3 + 2\sqrt{2}) \leq x \leq 1$ (the number $\frac{1}{8}(3 + 2\sqrt{2})$ being the larger of the two roots of the equation $x + \frac{1}{8} = \sqrt{x}$), and for infinitely many other functions of the family \mathcal{D} .

The value of c for given λ will be denoted by d_λ , and any function φ for which $d = d_\lambda$ will be called an approximating function, and represented by the notation $\varphi_\lambda(x)$.

4. Convergence for $m = \infty$. Let $d_{\lambda m}$ represent the maximum of $|f(x) - \varphi_{\lambda m}(x)|$ for $a \leq x \leq b$. If λ is held fast and m is allowed to become infinite, it will be shown that

$$(3) \quad \lim_{m \rightarrow \infty} d_{\lambda m} = d_\lambda.$$

As the functions $\varphi_{\lambda m}$, for fixed λ , are uniformly bounded* and equally continuous, they have one or more limit functions as m becomes infinite, uniformly approached by sequences suitably chosen from among them. When (3) has been established, it will follow that any one of these limit functions is a function ψ_λ solving the problem of § 3.

To prove (3), let ε be an arbitrary positive quantity. Suppose that $d_{\lambda m} \geq d_\lambda + \varepsilon$, for a specified value of m . Let x_0 be a value of x for which $|f - \varphi_{\lambda m}| = d_{\lambda m}$. By the assumed continuity of $f(x)$, there exists an $h_1 > 0$ such that

$$|f(x) - f(x_0)| \leq \frac{\varepsilon}{4}$$

for $|x - x_0| \leq h_1$. On the other hand,

$$|\varphi_{\lambda m}(x) - \varphi_{\lambda m}(x_0)| \leq \frac{\varepsilon}{4}$$

when $|x - x_0| \leq \varepsilon/(4\lambda)$. Let h be the smaller of the numbers h_1 and $\varepsilon/(4\lambda)$. This h is independent of m , though it of course depends on λ , which is understood to be held fast throughout the present discussion. For $|x - x_0| \leq h$,

$$|f(x) - \varphi_{\lambda m}(x)| \geq d_\lambda + \frac{\varepsilon}{2}.$$

* It is clear that $\mu \leq \varphi_{\lambda m} \leq M$, for all values of λ , m , and x , either directly from the definition of $\varphi_{\lambda m}$ or from the discussion by which its existence was established in § 2.

This relation then holds throughout an interval of length at least h , even if x_0 happens to coincide with a or b . Consequently

$$(4) \quad \int_a^b |f(x) - \varphi_{\lambda m}(x)|^m dx \geq h \left(d_\lambda + \frac{\varepsilon}{2} \right)^m.$$

But

$$(5) \quad \int_a^b |f(x) - \psi_\lambda(x)|^m dx \leq (b - a) d_\lambda^m.$$

As the left-hand member of (4) must be less than or equal to the left-hand member of (5), by the definition of $\varphi_{\lambda m}$, it must be that

$$h \left(d_\lambda + \frac{\varepsilon}{2} \right)^m \leq (b - a) d_\lambda^m,$$

a relation which ultimately ceases to hold when m becomes indefinitely large. So the assumed relation $d_{\lambda m} \geq d_\lambda + \varepsilon$ leads to a contradiction, as soon as m is sufficiently large, and it must be that $d_{\lambda m} < d_\lambda + \varepsilon$ for all values of m from a certain point on. As $d_{\lambda m} \geq d_\lambda$ throughout, from the definition of d_λ , the truth of (3) is established.

5. **Convergence for $\lambda = \infty$.** In the case of the approximating functions of § 3, it is almost immediately evident that*

$$\lim_{\lambda \rightarrow \infty} \psi_\lambda(x) = f(x),$$

uniformly for $a \leq x \leq b$. For $f(x)$, by reason of its uniform continuity, can be approximated with any desired accuracy by a continuous single-valued function whose graph is a broken line with a finite number of segments, and any function of this character belongs to the family Φ_λ as soon as λ is sufficiently large, while the definition of ψ_λ requires that its maximum error be not greater than that of any other function in Φ_λ .

The corresponding problem of convergence for the approximating functions

* When the approximating function is not unique, ψ_λ may be any approximating function corresponding to the value of λ in question.

$\varphi_{\lambda m}$, when m is held fast* and λ is allowed to become infinite, appears to be somewhat less trivial. Let $\omega(h)$ be the modulus of continuity of $f(x)$, the maximum of $|f(x') - f(x'')|$ for $|x' - x''| \leq h$. It will be shown that

$$\lim_{\lambda \rightarrow \infty} \varphi_{\lambda m}(x) = f(x).$$

uniformly for $a \leq x \leq b$, provided that a positive constant k and a constant $\alpha > 1/(m+1)$ exist so that $\omega(h) \leq kh^\alpha$ for $0 < h \leq b-a$. It may be assumed without loss of generality† that $\alpha < 1$.

For the purposes of the proof, it will be convenient to examine more closely the approximation of $f(x)$ by means of a broken-line function. Let h have any positive value $\leq b-a$. Let $q(x)$ be a broken-line function defined by the requirement that it shall coincide with $f(x)$ for $x = a, a+h, a+2h, \dots, a+nh$, where n is the greatest integer contained in $(b-a)/h$, and shall be equal to $f(b)$ for $x = a + (n+1)h$. In any one interval of length h , neither $f(x)$ nor $q(x)$ can change by more than $\omega(h)$, and consequently

$$|f(x) - q(x)| \leq 2\omega(h)$$

for $a \leq x \leq b$.

Let λ have any value $\geq k(b-a)^{\alpha-1}$. For this λ , let a number h be determined so that $\lambda h = kh^\alpha$, that is, let $h = (\lambda/k)^{1/(\alpha-1)}$. The slope of any segment of the graph of the function $q(x)$ defined above can not exceed $\omega(h)/h$, which is less than or equal to λ ; $q(x)$ belongs to the family Φ_λ , and $d_\lambda \leq 2\omega(h) \leq 2kh^\alpha$. By reason of the assumption that $\alpha < 1$, h will approach zero when λ is subsequently allowed to become infinite.

Let x_0 be a value of x for which $|f(x) - \varphi_{\lambda m}(x)|$ attains its maximum value $d_{\lambda m}$. Let β be a positive constant, presently to be specified. For $|x - x_0| \leq h^\beta$,

$$|f(x) - f(x_0)| \leq \omega(h^\beta) \leq kh^{\alpha\beta}, \quad |\varphi_{\lambda m}(x) - \varphi_{\lambda m}(x_0)| \leq \lambda h^\beta = kh^{\alpha+1-\beta-1},$$

and consequently

$$(6) \quad |f(x) - \varphi_{\lambda m}(x)| \geq d_{\lambda m} - kh^{\alpha\beta} - kh^{\alpha+1-\beta-1}.$$

* If $m < 1$, $\varphi_{\lambda m}$ is understood to represent an arbitrary one of the approximating functions, in case the determination is not unique. With this understanding, the conclusion holds for any value of $m > 0$.

† If $\alpha = 1$, $f(x)$ itself satisfies a Lipschitz condition, and is identical with $\varphi_{\lambda m}(x)$ for all values of m , as soon as λ is sufficiently large; if $\alpha > 1$, $f(x)$ is constant, and the problem becomes still more trivial.

Let it be assumed for the moment that the right-hand member here is positive or zero; the contrary hypothesis will be considered later. If β is held fast, h^β will certainly be less than $b - a$, at any rate for values of λ that are sufficiently large, and (6) will hold throughout an interval of length at least h^β , wherever x_0 may be situated. Then

$$\int_a^b |f(x) - \varphi_{\lambda m}(x)|^m dx \geq h^\beta [d_{\lambda m} - kh^{\alpha\beta} - kh^{\alpha+\beta-1}]^m.$$

Since $d_\lambda \leq 2kh^\alpha$,

$$\int_a^b |f(x) - \psi_\lambda(x)|^m dx \leq (b - a)(2kh^\alpha)^m.$$

The defining property of $\varphi_{\lambda m}$ then requires that

$$h^\beta [d_{\lambda m} - kh^{\alpha\beta} - kh^{\alpha+\beta-1}]^m \leq (b - a)(2kh^\alpha)^m,$$

$$d_{\lambda m} \leq kh^{\alpha\beta} + kh^{\alpha+\beta-1} + 2k \cdot \sqrt[m]{b - a} \cdot h^{\alpha-\beta/m}.$$

The last relation has been obtained on the assumption that $d_{\lambda m} \geq kh^{\alpha\beta} + kh^{\alpha+\beta-1}$, but clearly holds in the contrary case as well. The hypothesis of the theorem to be proved specifies that $\alpha > 1/(m + 1)$; if β is taken equal to $m/(m + 1)$, all three of the exponents on the right are positive, the entire right-hand member approaches zero as λ becomes infinite, and the uniform convergence of $\varphi_{\lambda m}(x)$ to the value of $f(x)$ is proved.

It is perhaps superfluous here to attempt further refinement of the sufficient condition that has been obtained for convergence, as it seems unlikely that the present method would ever yield anything even remotely approaching a necessary condition.

6. A property of the approximating functions $\varphi_{\lambda m}$. In conclusion, a property will be pointed out which in some degree, though by no means completely, characterizes the functions $\varphi_{\lambda m}$. Roughly expressed, the observation is that $\varphi_{\lambda m}$, in order to hold its title of approximating function, must make the utmost possible use of such freedom of oscillation as is left to it by the Lipschitz condition defining the family \mathcal{O}_λ , wherever it does not attain coincidence with $f(x)$.

To begin with a simple case, suppose $m = 2$, and suppose, if possible, that there is a constant $\lambda' < \lambda$, such that the quotient $|\varphi_{\lambda_2}(x_2) - \varphi_{\lambda_2}(x_1)|/|x_2 - x_1|$ is less than or equal to λ' throughout (a, b) . Let $\theta(x)$ be an arbitrary function for which the corresponding quotient never exceeds unity. Then, if r is a sufficiently small quantity, $\varphi_{\lambda_2}(x) + r\theta(x)$ belongs to \mathcal{D}_{λ} , and for $r \neq 0$ can not give so good an approximation as $\varphi_{\lambda_2}(x)$, according to the integral of the square of the error. Let

$$\delta(r) = \int_a^b [f(x) - \varphi_{\lambda_2}(x) - r\theta(x)]^2 dx.$$

Then

$$\frac{d}{dr} \delta(r) = \delta'(r) = -2 \int_a^b [f(x) - \varphi_{\lambda_2}(x) - r\theta(x)] \theta(x) dx.$$

Since $\delta(r)$ must have a minimum for $r = 0$, it must be that $\delta'(0) = 0$, that is,

$$\int_a^b [f(x) - \varphi_{\lambda_2}(x)] \theta(x) dx = 0.$$

This relation must hold for every function $\theta(x)$ which satisfies a Lipschitz condition with coefficient 1. But a θ can be chosen to violate it, if there is a point in (a, b) at which $f - \varphi_{\lambda_2} \neq 0$. So the approximating function $\varphi_{\lambda_2}(x)$ can not satisfy a Lipschitz condition with coefficient smaller than λ , unless it is identical with $f(x)$ throughout.

More generally, it would be easy to show, by an argument similar in character though different to some extent in detail, that if m has any value > 0 , if x_0 is any value of x in (a, b) for which $f - \varphi_{\lambda_m} \neq 0$, and if (a', b') is any interval containing x_0 as interior or end point, the quotient $|\varphi_{\lambda_m}(x_2) - \varphi_{\lambda_m}(x_1)|/|x_2 - x_1|$ must take on values arbitrarily near to λ in (a', b') . The underlying fact, however, appears to be that contained in the following somewhat less elementary theorem:

Let E_1 be the set of points in (a, b) where $\varphi_{\lambda_m}(x) = f(x)$, E_2 the set where φ_{λ_m} has a derivative equal to $\pm \lambda$, and E_0 the set where neither of these conditions is satisfied. Then the set E_0 is of measure zero.

As φ_{λ_m} is absolutely continuous,* it is known at the outset that it has a derivative almost everywhere. Assume that the theorem is not true, that is,

* Cf., e. g., de la Vallée Poussin, *Intégrales de Lebesgue*, Paris, 1916, pp. 76 ff.

that $mE_0 > 0$. Let $\varphi'_{\lambda m}$ be equal to the derivative of $\varphi_{\lambda m}$ where the derivative exists, and equal to zero elsewhere. There must be a positive number η such that $mE_3 > 0$, if E_3 consists of those points of E_0 at which $|f - \varphi_{\lambda m}| > \eta$. Let E_3' and E_3'' be the parts of E_3 where $f - \varphi_{\lambda m} > \eta$ and $f - \varphi_{\lambda m} < -\eta$ respectively. At least one of these sets must have a measure greater than zero, and it may be assumed for definiteness that $mE_3' > 0$. The set of all points where $f - \varphi_{\lambda m} > \eta$ is an open set (since the difference in question is a continuous function of x), and consists of a finite number or an enumerable infinity of open intervals. There must be at least one of these intervals which by itself contains a part of E_3' of measure greater than zero. Let (a', b') be such an interval, and let E_4 be the part of E_3' contained in (a', b') . Let $\theta'(x)$ be a function equal to $\lambda - |\varphi'_{\lambda m}(x)|$ (and therefore positive) at the points of E_4 , and equal to zero everywhere else, and let

$$\theta(x) = \int_a^x \theta'(x) dx,$$

the integral being taken in the sense of Lebesgue, if necessary. Then $\theta(x)$ is identically zero for $a \leq x \leq a'$, increases to a positive value A for $x = b'$ (since it is the integral of a function which is never negative, and which is actually positive throughout a set of positive measure between a' and b') and is constantly equal to A for $b' \leq x \leq b$. Being continuous, θ must take on the value $\frac{1}{2}A$ for at least one point $x = c'$ between a' and b' . Let

$$\theta_1'(x) = \theta'(x) \text{ for } a \leq x \leq c', \quad \theta_1'(x) = -\theta'(x) \text{ for } c' < x \leq b,$$

and let

$$\theta_1(x) = \int_a^x \theta_1'(x) dx.$$

The new function $\theta_1(x)$, likewise continuous, is identically zero for $a \leq x \leq a'$ and for $b' \leq x \leq b$, is never negative, and is actually positive for $x = c'$, and hence throughout some interval contained between a' and b' .

The function $\varphi_{\lambda m}(x) + r\theta_1(x)$ remains in the family Φ_λ for all values of r belonging to the interval $0 \leq r \leq 1$, because

$$|\varphi'_{\lambda m}(x) + r\theta_1'(x)| \leq |\varphi'_{\lambda m}(x)| + r|\theta_1'(x)| \leq \lambda,$$

and

$$\begin{aligned}
 & |[\varphi_{\lambda m}(x_2) + r\theta_1(x_2)] - [\varphi_{\lambda m}(x_1) + r\theta_1(x_1)]| \\
 &= \left| \int_{x_1}^{x_2} (\varphi'_{\lambda m} + r\theta'_1) dx \right| \leq \lambda |x_2 - x_1|.
 \end{aligned}$$

At the same time, $|\theta'_1| \leq \lambda$, and $|\theta_1| \leq \lambda(b-a)$, throughout (a, b) . If r is chosen to satisfy the requirements

$$0 < r \leq 1, \quad r \leq \frac{\eta}{\lambda(b-a)},$$

it is certain that $f - \varphi_{\lambda m} - r\theta_1$ is identical with $f - \varphi_{\lambda m}$ for $a \leq x \leq a'$ and for $b' \leq x \leq b$, that

$$0 < f - \varphi_{\lambda m} - r\theta_1 \leq f - \varphi_{\lambda m}$$

throughout the interior of (a', b') , and that the last relation is an actual inequality at the point $x = c'$ and throughout an interval containing this point. This means that the function $\varphi_{\lambda m} + r\theta_1$, belonging to \mathcal{O}_λ , gives a smaller value to the integral \mathcal{J} than the function $\varphi_{\lambda m}$, and the contradiction which was needed to establish the theorem has been obtained.

UNIVERSITY OF MINNESOTA,
MINNEAPOLIS, MINN.
