A GENERAL MEAN-VALUE THEOREM*

BY

D. V. WIDDER

In a paper published in 1906†, Professor G. D. Birkhoff treated the mean-value and remainder theorems belonging to polynomial interpolation, in which the linear differential operator $u^{(n)}$ played a particular rôle. It is natural to expect that a generalization of many of the ideas of that paper may apply to the general linear differential operator of order n, and the author is attempting such a program. This generalization throws fundamentally new light on the theory of trigonometric interpolation.

A very elegant paper by G. Pólya‡ has just appeared treating mean-value theorems for the general operator in a restricted interval. It is the special aim of the present paper to develop a general mean-value theorem, and to show how it can be specialized to obtain Pólya's results.

We consider a linear differential expression of order n,

$$Lu \equiv u^{(n)}(x) + l_1(x) u^{(n-1)}(x) + \cdots + l_n(x) u(x),$$

where $l_1(x)$, $l_2(x)$, $\cdots$, $l_n(x)$ are continuous functions, and $u(x)$ is continuous with its first $(n-1)$ derivatives, the nth derivative being piecewise continuous. All functions concerned are real. It is the purpose of this paper to obtain a necessary and sufficient condition for the change of sign of Lu in an interval in which u vanishes $(n+1)$ times.

More generally, the $(n+1)$ conditions implied in the vanishing of u may be replaced by an equal number of conditions involving also the derivatives of u. Let x_0, x_1, $\cdots$, x_n be points of the closed interval (a, b), which points need not be all distinct, and let k_0, k_1, $\cdots$, k_n be zero or positive integers not greater than $n-1$. We take then as $n+1$ conditions on u the relations

$$(A) \qquad u^{(k_i)}(x_i) = 0 \qquad (i = 0, 1, 2, \cdots, n).$$

* Presented to the Society, May 3, 1924.

† G. D. Birkhoff, *General mean-value and remainder theorems,* these Transactions, vol. 7 (1906), pp. 107–136. See also Bulletin of the American Mathematical Society, vol. 28 (1922), p. 5.

‡ G. Pólya, *On the mean-value theorem corresponding to a given linear homogeneous differential equation,* these Transactions, vol. 24 (1922), pp. 312–324.

Here $u^{(k_i)}(x)$ denotes the k_ith derivative of u, and $u^{(0)}(x)$ is the same as $u(x)$. We assume that no two of these equations are identical.

Now let $u_1, u_2, \cdots, u_n$ be n linearly independent solutions of the homogeneous equation

$$Lu = 0. \tag{1}$$

For definiteness take them as the principal solutions for the point a; that is, solutions satisfying the conditions

$$\begin{array}{lllll} u_1(a) = 0, & u_1'(a) = 0, & \cdots, & u_1^{(n-2)}(a) = 0, & u_1^{(n-1)}(a) = 1, \\ u_2(a) = 0, & u_2'(a) = 0, & \cdots, & u_2^{(n-2)}(a) = 1, & u_2^{(n-1)}(a) = 0, \\ \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot \\ u_n(a) = 1, & u_n'(a) = 0, & \cdots, & u_n^{(n-2)}(a) = 0, & u_n^{(n-1)}(a) = 0. \end{array}$$

We consider also the non-homogeneous equation

$$Lu = \varphi(x), \tag{2}$$

where $\varphi(x)$ is piece-wise continuous in (a, b). The general solution of (2) is now obtained by Cauchy's method. Determine a solution of (1) which together with its first $n-2$ derivatives vanishes at a point t of (a, b), while the $(n-1)$th derivative has the value unity at that point. Denote the function by $g(x, t)$. It satisfies the n conditions

$$g(t,t) = 0, \quad g'(t,t) = 0, \quad \cdots, \quad g^{(n-2)}(t,t) = 0, \quad g^{(n-1)}(t,t) = 1. \tag{3}$$

Here the differentiation is with respect to the first argument.

It is known that the general solution of (2) may be written in the form

$$u(x) = c_1 u_1 + c_2 u_2 + \cdots + c_n u_n + \tfrac{1}{2}\int_a^b \pm g(x,t)\varphi(t)\,dt.$$

Here $c_1, c_2, \cdots, c_n$ are arbitrary constants, and the sign before $g(x,t)$ is to be taken positive if $t < x$ and negative if $t > x$.* Considered as a function of t, $g(x, t)$ satisfies the equation adjoint to (1),

$$M(v) \equiv (-1)^n \frac{d^n v}{dx^n} + (-1)^{(n-1)} \frac{d^{n-1}}{dx^{n-1}}(l_1 v) + \cdots - \frac{d}{dx}(l_{n-1} v) + l_n v = 0.$$

* See for example D. A. Westfall, Dissertation, p. 16.

If now we express the fact that $u(x)$ satisfies the $n+1$ conditions (A), we obtain $n+1$ equations

$$(4)\quad c_1 u_1^{(k_i)}(x_i) + c_2 u_2^{(k_i)}(x_i) + \cdots + c_n u_n^{(k_i)}(x_i) + \tfrac{1}{2}\int_a^b \pm g^{(k_i)}(x_i, t)\varphi(t)\,dt = 0$$

$$(i = 0, 1, 2, \cdots, n).$$

Eliminating $c_1, c_2, \cdots, c_n$ from these equations, there results an equation of the form

$$(5)\qquad \int_a^b \Delta(t)\varphi(t)\,dt = 0,$$

where $\Delta(t)$ is the determinant

$$|\pm\tfrac{1}{2}g^{(k_i)}(x_i, t)\, u_1^{(k_i)}(x_i)\, u_2^{(k_i)}(x_i) \cdots u_n^{(k_i)}(x_i)| \qquad (i = 0, 1, \cdots, n).$$

Denote the cofactors of the elements of the first column by $\Delta_0, \Delta_1, \cdots, \Delta_n$, so that $\Delta(t)$ takes the form

$$\Delta(t) = \sum_{i=0}^{n} \pm\tfrac{1}{2} g^{(k_i)}(x_i, t)\Delta_i.$$

It is evident that $\Delta(t)$ depends in no way on the choice of the linearly independent solutions $u_1, u_2, \cdots, u_n$, but merely on the position of the points $x_0, x_1, \cdots, x_n$.

Now let us suppose that $\Delta(t)$ is not identically zero. Then the function u satisfying the conditions (A) can not be a solution of (1) unless it is identically zero; for a necessary and sufficient condition that there exist a solution of (1) not identically zero and satisfying the conditions (A) is precisely that

$$\Delta_i = 0 \qquad (i = 0, 1, \cdots, n).$$

Then $\varphi(t)$ is not identically zero, and we have at once from (5) a sufficient condition that Lu change sign in the interval (a, b), namely that $\Delta(t)$ should be a function of one sign in that interval. We may in particular take a and b as the two points of the set $x_0, x_1, \cdots, x_n$ which are farthest apart, and thus be assured that the change of sign of Lu occurs between these two points.

The condition is also necessary. Suppose that any function u with the required degree of continuity which is not identically zero and which satisfies the conditions (A) is such that Lu changes sign between the extreme points. It is desired to show that $\Delta(t)$ is a function of one sign not identically zero.

We note first that the Δ_i are not all zero; for if they were, there would be a solution u of (1) satisfying the conditions (A). This is impossible since Lu must change sign by hypothesis. In order to prove that $\Delta(t)$ does not vanish identically we must investigate its structure more closely.

If we denote by $v_1, v_2, \cdots, v_n$ the solutions adjoint to $u_1, u_2, \cdots, u_n$, we may write $g(x, t)$ as follows:*

$$g(x, t) = \sum_{i=1}^{n} u_i(x)\, v_i(t).$$

$\Delta(t)$ then takes the form

$$\Delta(t) = \tfrac{1}{2}\Big[v_1(t) \sum_{i=0}^{n} \pm u_1^{(k_i)}(x_i)\, \Delta_i + v_2(t) \sum_{i=0}^{n} \pm u_2^{(k_i)}(x_i)\, \Delta_i + \cdots$$
$$\cdots + v_n(t) \sum_{i=0}^{n} \pm u_n^{(k_i)}(x_i)\, \Delta_i\Big].$$

If $\Delta(t)$ were identically zero, each summation in the above expression would be zero, since $v_1, v_2, \cdots, v_n$ are linearly independent. Now by taking t in all possible positions with respect to the points $x_0, x_1, \cdots, x_n$, various combinations of signs in each summation are obtained.

It would follow then that

$$u_j^{(k_i)}(x_i)\, \Delta_i = 0 \qquad (i = 0, 1, \cdots, n;\ j = 1, 2, \cdots, n).$$

Not all the Δ_i are zero. Suppose that $\Delta_m \neq 0$. Then it would follow that

$$u_j^{(k_m)}(x_m) = 0 \qquad (j = 1, 2, \cdots, n).$$

The Wronskian of $u_1, u_2, \cdots, u_n$ would vanish at the point x_m, contrary to the assumption that $u_1, u_2, \cdots, u_n$ are linearly independent. $\Delta(t)$ can not therefore vanish identically.

Suppose now that $\Delta(t)$ changes sign between the extreme points. It is then possible to choose a continuous† function of one sign $\bar{\varphi}(t)$ such that

$$(6) \qquad \int_a^b \bar{\varphi}(t)\, \Delta(t)\, dt = 0.$$

* See for example, Darboux, *Théorie des Surfaces*, vol. 2, p. 106.

† Indeed $\bar{\varphi}$ may be a piece-wise continuous function made up of straight lines parallel to the x-axis.

Now consider the equations (4) in which $\varphi(t)$ is replaced by $\bar{\varphi}(t)$. From these $n+1$ equations, pick out that set of n equations which has Δ_m for its determinant. Since $\Delta_m \neq 0$, the set of equations has a unique solution in $c_1, c_2, \cdots, c_n$. If we substitute this solution in the equation

$$\bar{u}^{(k_m)}(x) = c_1 u_1(x) + \cdots + c_n u_n(x) + \tfrac{1}{2}\int_a^b \pm g^{(k_m)}(x, t)\,\bar{\varphi}(t)\,dt,$$

we obtain an equation of the form

$$\bar{u}^{(k_m)}(x) = \int_a^b G^{(k_m)}(x, t)\,\bar{\varphi}(t)\,dt,$$

where $G(x, t)$ may be identified with the Green's function corresponding to the boundary conditions (A).*

It is seen that

$$G^{(k_m)}(x_m, t) = \Delta(t).$$

It follows from (6) that

$$\bar{u}^{(k_m)}(x_m) = 0.$$

By its very definition $\bar{u}(x)$ is seen to satisfy the remainder of the conditions (A). $L\bar{u}$ must therefore change sign. But $L\bar{u}$ is $\bar{\varphi}$, a function which does not change sign. We have thus completed the proof of the following

THEOREM I. *A necessary and sufficient condition that Lu change sign in an interval in which u (having the required degree of continuity† and not identically zero) satisfies the conditions (A) is that $\Delta(t)$ be a function of one sign not identically zero in that interval.*

G. Pólya obtains certain theorems concerning the vanishing of Lu. We may obtain these results from Theorem I.

With Pólya we say that the property W holds for the operator Lu in an open interval (a, b) if there exist solutions of (1), $h_1, h_2, \cdots, h_{n-1}$, such that the following functions do not vanish in (a, b):

$$W_1 = h_1, \qquad W_2 = W(h_1, h_2) = \begin{vmatrix} h_1 & h_2 \\ h_1' & h_2' \end{vmatrix}, \quad \cdots$$

$$\cdots, \qquad W_{n-1} = W(h_1, \cdots, h_{n-1}) = \begin{vmatrix} h_1 & h_2 & \cdots & h_{n-1} \\ h_1' & h_2' & \cdots & h_{n-1}' \\ \cdot & \cdot & \cdot & \cdot \\ h_1^{(n-2)} & h_2^{(n-2)} & \cdots & h_{n-1}^{(n-2)} \end{vmatrix}.$$

* C. E. Wilder, these Transactions, vol. 18 (1917), p. 416.

† The restrictions on the nth derivative of u might be made lighter as is done in Birkhoff's paper, loc. cit.

Pólya considers the special case of the conditions (A) which involves the vanishing of $u(x)$ at points of the interval that are distinct or coincident. Consider r points

$$x_1 < x_2 < \cdots < x_r, \qquad r \leqq n+1.$$

Suppose that $u(x)$ vanishes m_i times at a point x_i $(m_i \leqq n-1)$:

$$(B) \qquad \begin{gathered} u(x_i) = u'(x_i) = \cdots = u^{(m_i-1)}(x_i) = 0, \quad u^{(m_i)}(x_i) \neq 0, \quad i = 1, 2, \cdots, r; \\ \sum_{i=1}^{r} m_i = n+1. \end{gathered}$$

If $x_1, x_2, \cdots, x_r$ lie in the interval in which the property W holds, $\Delta(t)$ is a function of one sign. To prove this we investigate the structure of $\Delta(t)$. In each of the $r-1$ intervals, $\Delta(t)$ is a solution of the adjoint equation, continuous with its first n derivatives. It is only at a point x_i that a discontinuity may occur. Such a discontinuity is caused by a change in the ambiguous sign before $g^{(k_i)}(x_i, t)$ as t passes over x_i. However, it is only when $g^{(k_i)}(x_i, x_i)$ is not zero that such a discontinuity is introduced. We can now show that at a point x_i where $u(x)$ vanishes m_i times $\Delta(t)$ is continuous with its first $n-m_i-1$ derivatives.

It is known* that

$$\frac{\partial^{\mu}}{\partial x^{\mu}} \frac{\partial^{\nu}}{\partial t^{\nu}} g(x, t) \bigg|_{t=x} = \sum_{i=1}^{n} u_i^{(\mu)}(x) v_i^{(\nu)}(t) \bigg|_{t=x} = 0, \qquad \mu + \nu < n-1,$$

$$= (-1)^{\nu}, \quad \mu + \nu = n-1.$$

At x_i, μ may equal $m_i - 1$, so that

$$\frac{\partial^{\mu}}{\partial x^{\mu}} \frac{\partial^{\nu}}{\partial t^{\nu}} g(x_i, x_i) = 0, \qquad \mu < m_i, \qquad \nu < n - m_i.$$

It follows then that $\Delta(t)$ is continuous at x_i with its first $n-m_i-1$ derivatives.

Now if $t < x_i$, $i = 1, 2, \cdots, r$, then all the ambiguous signs in $\Delta(t)$ are positive; if $t > x_i$, $i = 1, 2, \cdots, r$, they are all negative. In either case $\Delta(t)$ is identically zero for the interval considered, since

$$\sum_{i=0}^{n} u_j^{(k_i)}(x_i)\, \Delta_i = 0 \qquad (j = 1, 2, \cdots, n),$$

* See Schlesinger, *Lineare Differential-Gleichungen*, vol. 1, p. 63.

the expression on the left being a determinant with two columns equal. Now since $\Delta(t)$ is continuous with its first $n-m_1-1$ derivatives at x_1, it follows that $\Delta(t)$ has $n-m_1$ zeros at x_1. It has $n-m_r$ zeros at x_r.

Suppose first that

$$m_1 = m_2 = \cdots = m_{n+1} = 1, \qquad x_r = x_{n+1}.$$

Then $\Delta(t)$ is continuous throughout with its first $n-2$ derivatives and has $n-1$ zeros in each of the points x_1 and x_r. Suppose it were not a function of one sign. It would have in all at least $2n-1$ zeros in the closed interval (x_1, x_r).

Now recall that $M(v)$ can be "factored" as follows:

$$M(v) \equiv \frac{(-1)^n}{W_1} \frac{d}{dx} \frac{W_1^2}{W_2 W_0} \frac{d}{dx} \frac{W_2^2}{W_3 W_1} \frac{d}{dx} \cdots \frac{d}{dx} \frac{W_{n-1}^2}{W_n W_{n-2}} \frac{d}{dx} \frac{W_n}{W_{n-1}} v,$$

where $W_0 = 1$.* Each of the quantities $W_0, W_1, \cdots, W_n$ does not vanish in the interval (x_1, x_{n+1}), so that we may apply Rolle's theorem. If $\Delta(t)$ vanished $2n-1$ times in (x_1, x_{n+1}), the function

$$\psi(t) = \frac{W_1^2}{W_2 W_0} \frac{d}{dx} \frac{W_2^2}{W_1 W_3} \frac{d}{dx} \cdots \frac{d}{dx} \frac{W_n}{W_{n-1}} \Delta(t)$$

would change sign at least n times *inside* the interval. But $M(\Delta(t))$ is identically zero, so that $\psi(t)$ is constant in any interval in which it is continuous. Hence $\psi(t)$ can change sign only at the points $x_2, x_3, \cdots, x_n$, where it is discontinuous. There are only $n-1$ such points, so that it must be concluded that $\Delta(t)$ is a function of one sign in (x_1, x_{n+1}).

In order to obtain the proof in the general case we shall need the following

LEMMA. *If a function $f(x)$ is continuous in an interval in which $f'(x)$ is continuous except for l finite jumps, and if $f'(x)$ can have at most N zeros in the interval, $f(x)$ can have at most $N+l+1$ zeros there.*

Proof. If $f(x)$ had $N+l+2$ zeros, $f'(x)$ would have $N+l+1$ zeros and discontinuous changes of sign. At most l of these can be discontinuous changes of sign, and $f'(x)$ would have $N+1$ zeros contrary to hypothesis.

Now denote by s_i the number of integers $m_2, m_3, \cdots, m_{r-1}$, which are equal to i. Then

(7) $$s_1 + s_2 + \cdots = r-2.$$

* See Schlesinger, loc. cit., p. 58.

Define a function $\bar{\Delta}^{(k)}(t)$ by the equation

$$\bar{\Delta}^{(k)}(t) = \frac{W_{n-k}^2}{W_{n-k-1}\,W_{n-k+1}}\,\frac{d}{dx}\,\frac{W_{n-k+1}^2}{W_{n-k}\,W_{n-k+2}}\,\frac{d}{dx}\cdots\frac{d}{dx}\,\frac{W_n}{W_{n-1}}\,\Delta(t).$$

Now apply the lemma to $\bar{\Delta}^{(n-2)}$ in each of the $r-s_1-1$ intervals in which it is continuous. We may evidently treat all these intervals simultaneously and take l equal to the total number of discontinuities of $\bar{\Delta}^{(n-1)}$, viz., $r-2$. N must be zero since $\bar{\Delta}^{(n-1)}$ is constant where it is continuous.* We conclude that $\bar{\Delta}^{(n-2)}$ can vanish at most $r-1$ times.

Now apply the lemma to $\bar{\Delta}^{(n-3)}$. Here $N=r-1$ and $l=r-s_1-2$. Then $\bar{\Delta}^{(n-3)}$ can have at most $2(r-1)-s_1$ zeros. $\bar{\Delta}^{(n-4)}$ can have at most $3(r-1)-2s_1-s_2$. Proceeding in this way we see that $\Delta(t)$ has at most

$$(n-1)(r-1)-(n-2)s_1-(n-3)s_2-\cdots$$

zeros. Now $\Delta(t)$ has

$$n-m_1+n-m_r = n-1+s_1+2s_2+3s_3+\cdots$$

zeros at the end points x_1 and x_r, and this number is precisely equal to the maximum number of zeros $\Delta(t)$ can have, since by virtue of (7)

$$(n-1)(r-1)-(n-2)s_1-(n-3)s_2-\cdots = n-1+s_1+2s_2+\cdots.$$

The proof is thus complete that $\Delta(t)$ can not change sign.

Before proceeding to the converse of this theorem, let us draw several further inferences.

THEOREM II. *In an interval in which the property W holds, the coefficient of* $g^{(m_i-1)}(x_i,t)$ *in* $\Delta(t)$ *can not vanish,* $i=1,2,\cdots,r$.

For if it did $\Delta(t)$ would be continuous with its first $n-m_i$ derivatives at x_i, and by means of the lemma a contradiction would be reached as before.

COROLLARY. *No solution of equation* (1) *can vanish n times in an interval in which the property W holds unless it is identically zero.*

For the vanishing of the coefficient Δ_i of Theorem II is precisely the condition that there exist a solution of (1) not identically zero passing through the n points (some of which may be coincident) involved in Δ_i.

* $\bar{\Delta}^{(n-1)}(t)$ is not identically zero in any interval between x_1 and x_r.

These n points were arbitrary in the interval so that the corollary is established.

Now let us show that if $\Delta(t)$ is a function of one sign in an interval $a \leqq x < b$ for every set of conditions (B) in this interval, then the property W holds in the interval $a < x < b$. We prove first that no solution of (1) can vanish n times in the interval $a \leqq x < b$. For suppose there were such a solution u. Let x_r be the point of vanishing nearest b, and let x' be a point between x_r and b. Now determine a solution w of the differential system

$$Lw = 1,$$
$$w(x') = w'(x') = \cdots = w^{(n-1)}(x') = 0.$$

Form the function $\bar{u}(x)$ which is

$$\begin{aligned} &u(x) + Mw(x), \qquad && x' \leqq x < b, \\ &u(x), && a \leqq x < x', \end{aligned}$$

where M is a constant to be determined. We wish to show that M can be so determined that $\bar{u}(x)$ vanishes $n+1$ times in $a \leqq x < b$. Now $u(x') \neq 0$ since $x_r < x'$, and hence it follows that $\bar{u}(x')$ has the sign of $u(x')$. Choose a point x'' between x' and b for which $w(x'')$ is not zero. Such a point exists since $w(x) \equiv\!\!\!\!/\ 0$ in any interval. Now choose M so that $\bar{u}(x'')$ will have a sign opposite to that of $\bar{u}(x')$; $\bar{u}$ will then vanish between x' and x''. $\bar{u}$ has the required degree of continuity to apply the mean-value theorem, and $\Delta(t)$ is a function of one sign. Hence $L\bar{u}$ must change sign. But $L\bar{u}$ is equal to zero in the interval $a < x < x'$ and to M in the interval $x' < x < b$, and does not change sign. The contradiction shows that u can not vanish n times in $a \leqq x < b$.

But if no solution of (1) vanishes n times in $a \leqq x < b$, it is a simple matter to show that the property W holds in $a < x < b$. For the principal solutions for the point a, $u_1, u_2, \cdots, u_n$, are suitable functions. The Wronskian $W_k = W(u_1, u_2, \cdots, u_k)$ does not vanish in $a < x < b$. For suppose it vanished at a point c of that interval. Then a function

$$c_1 u_1 + c_2 u_2 + \cdots + c_k u_k$$

could be determined not identically zero and having k zeros in the point c. But this function would have $n-k$ zeros in a, and a total of n zeros in $a \leqq x < b$. This is impossible. We may now state the following

THEOREM III. *A necessary and sufficient condition that the vanishing of a function u (with the required degree of continuity and not identically zero) at $n+1$ arbitrary points of an interval (a, b) should imply the change of sign of Lu at an intermediate point is that the property W hold in (a, b).*

A simple example will suffice to show that Theorem I is stronger than Theorem III. Take

$$Lu = u'' + u,$$

$$x_0 = 0 < x_1 < x_2.$$

Then

$$\Delta(t) = \frac{1}{2}\begin{vmatrix} \sin t & 0 & 1 \\ \pm\sin(x_1 - t) & \sin x_1 & \cos x_1 \\ \sin(x_2 - t) & \sin x_2 & \cos x_2 \end{vmatrix},$$

$$\Delta(t) = \sin t \sin(x_1 - x_2), \qquad 0 < t < x_1,$$

$$= \sin x_1 \sin(t - x_2), \qquad x_1 < t < x_2.$$

Suppose now that $x_1 < \pi$ and $x_2 - x_1 < \pi$.

Then

$$\Delta(t) < 0, \qquad 0 < t < x_2.$$

$\Delta(t)$ is a function of one sign in the interval $(0, x_2)$ which may clearly be of length greater than π. (In fact it may be as near to 2π as we like.) Yet the property W can not hold in any interval of length greater than π, in as much as some solution of (1) will vanish twice in such an interval. This example suggests possible generalizations of Pólya's results.

It should be pointed out that Theorem I might easily be made to apply to the most general linear boundary conditions.

HARVARD UNIVERSITY,
CAMBRIDGE, MASS.