THE STIELTJES MOMENT PROBLEM FOR FUNCTIONS OF BOUNDED VARIATION

R. P. BOAS, JR.*

1. Introduction. We shall establish the following theorem, which at first sight appears quite unexpected:

Theorem 1. Any sequence \( \{\mu_n\} \) of real numbers can be represented in the form

\[
\mu_n = \int_0^\infty t^n \alpha(t) \, dt, \quad n = 0, 1, 2, \ldots,
\]

(1.1)

\[
\int_0^\infty |\alpha(t)| < \infty.
\]

The problem of determining necessary and sufficient conditions for a sequence of numbers \( \{\mu_n\} \) to have the form

(1.2) \( \mu_n = \int_0^\infty t^n \alpha(t) \, dt, \quad \alpha(t) \) non-decreasing, \( n = 0, 1, 2, \ldots \),

was set and solved by T. J. Stieltjes. It would be natural to attempt to generalize the problem by requiring merely that \( \alpha(t) \) should be a function of bounded variation on \((0, \infty)\); but the generalized problem has, as Theorem 1 shows, a trivial solution.

To establish Theorem 1, we shall exhibit an arbitrary real sequence \( \{\mu_n\} \) as the difference of two sequences \( \{\lambda_n\} \) and \( \{\nu_n\} \), each of the form (1.2).† The construction will also lead to the result that any sequence \( \{\mu_n\} \) of positive numbers of sufficiently rapid growth has the form (1.2); it is sufficient, for example, that

(1.3) \( \mu_0 \geq 1, \quad \mu_n \geq (n \mu_{n-1})^n, \quad n \geq 1. \)

A specimen sequence satisfying (1.3) is \( \mu_0 = 1, \mu_n = n^n, (n = 1, 2, \ldots) \).

As an application‡ of Theorem 1, it will be shown that

* National Research Fellow.

† Added in proof: Other proofs of Theorem 1 have been given by G. Pólya (Sur l'indétermination d'un problème voisin du problème des moments, Comptes Rendus de l'Académie des Sciences, Paris, vol. 207 (1938), pp. 708–711). Pólya points out that a theorem of which Theorem 1 is an immediate consequence was proved by É. Borel in 1894.

\[ f(x) = \int_0^\infty x(t) \alpha(t), \]

with

\[ \int_0^\infty t^n |\alpha(t)| < \infty, \quad n = 1, 2, \ldots, \]

is not the general linear functional on any very interesting space of functions \( x = x(t) \), containing an infinite number of the functions \( t^n \), \( (n = 1, 2, \ldots) \) (see §4 for a precise statement). Other negative results of this character have been obtained by J. W. Tukey and the author;* the reader is referred to their paper for a discussion of the significance of such results.

2. Proof of Theorem 1. We use the notation

\[
\begin{bmatrix}
\mu_0 & \mu_1 & \cdots & \mu_n \\
\mu_1 & \mu_2 & \cdots & \mu_{n+1} \\
\vdots & \vdots & \ddots & \vdots \\
\mu_n & \mu_{n+1} & \cdots & \mu_{2n}
\end{bmatrix}, \quad n = 0, 1, 2, \ldots.
\]

Then a necessary and sufficient condition for \( \{\mu_n\} \) to have the form 1.2 is†

\[
(2.1) \quad [\mu_0 \mu_2 \cdots \mu_{2n}] \geq 0, \quad [\mu_1 \mu_3 \cdots \mu_{2n+1}] \geq 0, \quad n = 0, 1, 2, \ldots.
\]

We choose positive numbers \( \lambda_0, \lambda_1, \nu_0, \nu_1 \), so that \( \lambda_0 - \nu_0 = \mu_0, \lambda_1 - \nu_1 = \mu_1 \). We now proceed to define the sequences \( \{\lambda_n\}, \{\nu_n\} \) by induction. Suppose that

\[
(2.2) \quad \lambda_k - \nu_k = \mu_k
\]

for \( k = 0, 1, 2, \ldots, 2n-1 \), and that the determinants

\[
(2.3) \quad \begin{bmatrix}
\lambda_0 \lambda_2 & \cdots & \lambda_{2k} \\
\nu_0 \nu_2 & \cdots & \nu_{2k}
\end{bmatrix},
\begin{bmatrix}
\lambda_1 \lambda_3 & \cdots & \lambda_{2k+1} \\
\nu_1 \nu_3 & \cdots & \nu_{2k+1}
\end{bmatrix},
\]

are positive for \( k = 0, 1, 2, \ldots, n-1 \). We have (with undetermined \( \lambda_{2n} \))

\[
[\lambda_0 \lambda_2 \cdots \lambda_{2n}] = \lambda_{2n} [\lambda_0 \lambda_2 \cdots \lambda_{2n-2}] + P,
\]

where \( P \) is a polynomial in \( \lambda_0, \lambda_1, \cdots, \lambda_{2n-1} \); and there is a corre-

† See, for example, O. Perron, Die Lehre von den Kettenbrüchen, 1929, p. 410; cf. also M. Riesz, Sur le problème des moments, troisième note, Arkiv för Matematik, Astronomi och Fysik, vol. 17 (1922–1923), no. 16.
sponding relation for \([\nu_0\nu_2 \cdots \nu_{2n}]\). Since \([\lambda_0\lambda_2 \cdots \lambda_{2n-2}] > 0\), and
\([\nu_0\nu_2 \cdots \nu_{2n-2}] > 0\), we can choose \(\lambda_{2n}\) and \(\nu_{2n}\), so that \(\lambda_{2n} - \nu_{2n} = \mu_{2n}\),
and so large that \([\lambda_0\lambda_2 \cdots \lambda_{2n}] > 0\), \([\nu_0\nu_2 \cdots \nu_{2n}] > 0\). Similarly
we can then choose \(\lambda_{2n+1}\) and \(\nu_{2n+1}\), so that \(\lambda_{2n+1} - \nu_{2n+1} = \mu_{2n+1}\),
\([\lambda_0\lambda_2 \cdots \lambda_{2n+1}] > 0\), \([\nu_0\nu_3 \cdots \nu_{2n+1}] > 0\). This completes the induction:
we can find sequences \(\{\lambda_n\}, \{\nu_n\}\) such that for \(k = 0, 1, 2, \cdots\),
(2.2) is satisfied, and all the determinants (2.3) are positive. Then
\(\{\lambda_n\}\) and \(\{\nu_n\}\) satisfy (2.1), and consequently have the form (1.2),
so that \(\{\mu_n\}\) has the form (1.1).

3. Rapidly increasing sequences. We now prove the following theorem:

**Theorem 2.** If

\[
\mu_0 \geq 1, \quad \mu_n \geq (n\mu_{n-1})^n, \quad n = 1, 2, \cdots,
\]

then \(\{\mu_n\}\) has the form (1.2).

For the proof, we modify the construction of the sequence \(\{\lambda_n\}\) of
§2. We have, for \(n = 1, 2, \cdots\),

\[
[\mu_0\mu_2 \cdots \mu_{2n}] = \mu_{2n}[\mu_0\mu_2 \cdots \mu_{2n-2}] + \sum_{k=n}^{2n-1} \pm \mu_k D_k,
\]

where the \(D_k\) are \(n\)-rowed minors of \([\mu_0\mu_2 \cdots \mu_{2n}]\) and do not in­
volve \(\mu_{2n}\). Similarly, for \(n = 1, 2, \cdots\),

\[
[\mu_1\mu_3 \cdots \mu_{2n+1}] = \mu_{2n+1}[\mu_1\mu_3 \cdots \mu_{2n-1}] + \sum_{k=n+1}^{2n} \pm \mu_k D'_k,
\]

where the \(D'_k\) are \(n\)-rowed minors of \([\mu_1\mu_3 \cdots \mu_{2n+1}]\), not involving
\(\mu_{2n+1}\).

Suppose that for \(k \leq n - 1\), \((n = 1, 2, \cdots)\),

\[
[\mu_0\mu_2 \cdots \mu_{2k}] \geq 1, \quad [\mu_1\mu_3 \cdots \mu_{2k+1}] \geq 1.
\]

Assuming (3.1), we shall show that (3.4) is satisfied also for \(k = n\).

Clearly, \(\mu_m \geq 1\) for \(m = 1, 2, \cdots\). Hence we have

\[
\mu_m \geq (m\mu_{m-1})^m > 2(m/2)^{(m+4)/4 (m+2)/2} \mu_{m-1}, \quad m = 2, 3, \cdots.
\]

Therefore

\[
\mu_{2n} > 1 + n^{(n+2)/2} \mu_{2n-1}, \quad \mu_{2n+1} > 1 + n^{(n+2)/2} \mu_{2n}.
\]

Now, (3.1) implies in particular that \(\mu_{m+1} \geq \mu_m\), \((m \geq 1)\); hence the
elements of the determinants \(D_k\) do not exceed \(\mu_{2n-1}\), and the ele-
ments of the $D_k'$ do not exceed $\mu_{2n}$. Then by Hadamard’s theorem,*

$$|D_k| \leq \mu_{2n-1} n^{n/2}, \quad k = n, n + 1, \ldots, 2n - 1,$$

$$|D_k'| \leq \mu_{2n} n^{n/2}, \quad k = n + 1, n + 2, \ldots, 2n.$$

Therefore, using (3.2), (3.3), (3.4), (3.5), we obtain

$$[\mu_0, \mu_2, \ldots, \mu_{2n}] \geq \mu_{2n} - n \frac{n+1}{\mu_{2n-1}} > 1,$$

$$[\mu_1, \mu_3, \ldots, \mu_{2n+1}] \geq \mu_{2n+1} - n \frac{n+1}{\mu_{2n}} > 1.$$

Thus (3.4) holds for $k = n$ if it holds for $k < n$; but it holds for $k = 0$ by assumption, and consequently holds for all $k$; therefore $\{\mu_n\}$ has the form (1.2).

The moment problem (1.2) is said to be determined or undetermined according as the function $\alpha(t)$ is or is not unique (after being normalized by the conditions $\alpha(0) = 0$, $\alpha(t) = [\alpha(t+)+\alpha(t-)}/2$ for $t \neq 0$). A consequence of Theorem 2 is that the moment problem (1.2) is not only solvable for any sequence $\{\mu_n\}$ of sufficiently rapid growth, but is even undetermined. In fact, if $\{\mu_n\}$ satisfies (3.1) and if in addition $\mu_2 \geq (2\mu_1 + 2)^2$, we define a sequence $\{\nu_n\}$ by setting $\nu_1 = \mu_1 + 1$, $\nu_n = \mu_n$ for $n \neq 1$. Then $\{\nu_n\}$ satisfies (3.1); consequently for $n = 0, 1, 2, \ldots$,

$$\nu_{2n} = \int_0^\infty t^{2n} d\beta(t) = \int_0^\infty u^{n} d\beta(u^{1/2}) = \int_0^\infty u^n d\gamma(u),$$

say; while

$$\nu_{2n} = \mu_{2n} = \int_0^\infty t^{2n} d\alpha(t) = \int_0^\infty u^n d\delta(u),$$

where $\gamma(u)$ and $\delta(u)$ are normalized and non-decreasing. But $\gamma(u)$ and $\delta(u)$ are distinct, since

$$\nu_1 = \int_0^\infty u^{1/2} d\gamma(u) = 1 + \int_0^\infty u^{1/2} d\delta(u) = 1 + \mu_1.$$

Hence the moment problem for the sequence $\{\mu_{2n}\}$ is undetermined.

4. **Linear functionals.** We use the terminology of S. Banach’s book.† Let $R$ be a topological vector space of elements $x$, let $P$ be a

---


† *Théorie des Opérations Linéaires*, 1932.
space of elements $p$, and let $f_p(x)$ be a functional with domain $R$, defined for each $p$ in $P$. We say that a general linear functional in $R$ is $f_p(x)$, if the following conditions are satisfied:

(i) $f_p(x)$ is a linear functional for every $p \in P$.

(ii) Every linear functional $g(x)$ with domain $R$ is identically equal to some $f_p(x)$.

In the application to be made here, the elements of $P$ are the functions $p = p(t)$, of bounded variation on $(0, \infty)$, such that

$$
\int_0^\infty t^n |dp(t)| < \infty, \quad n = 1, 2, \ldots;
$$

the elements of $R$ are measurable functions $x = x(t)$, defined on $(0, \infty)$; and

$$
f_p(x) = \int_0^\infty x(t) dp(t),
$$

where the integral is a Lebesgue-Stieltjes integral. We have the following result:

**Theorem 3.** Let $R$ be a topological vector space with the following property:*  

(Q): If $x \in R$ and $a_n \to 0$, then $a_n x \to \Theta$.†

Then if $R$ contains an infinite number of functions $t^n$, ($n = 0, 1, 2, \ldots$), there is some $p \in P$ for which (4.1) is not a linear functional on $R$.

In particular, we see that, under the hypotheses of Theorem 3, (4.1) is not a general linear functional on $R$.

Suppose that (4.1) is, for every $p \in P$, a linear functional on a space $R$ with the specified properties. Let $S$ be the subspace composed of all finite linear combinations of the elements $t^n$ which are in $R$ (with the topology of $R$). If $f$ is an arbitrary distributive (that is, additive and homogeneous) functional with domain $S$, we define a sequence $\{\mu_n\}$ by setting $\mu_n = f(t^n)$ when $t^n \in R$, and $\mu_n = 0$ otherwise. By Theorem 1, there is a $p \in P$ such that

$$
\mu_n = \int_0^\infty t^n dp(t), \quad n = 0, 1, 2, \ldots.
$$

Since $f$ is distributive, we then have

---

* In particular, a space of type $F$ has this property.
† $\Theta$ denotes the zero element of $R$. 

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use
404

H. T. ENGSTROM

(4.2) \[ f(x) = \int_0^\infty x(t) \, dp(t), \quad x \in S. \]

Now (4.1) is a linear functional on \( R \), and consequently a linear functional on \( S \). Hence (4.2) states that every distributive functional on \( S \) is linear; but this is impossible unless \( S \) is finite-dimensional,* which it is not. This contradiction establishes the theorem.

NORTON, MASSACHUSETTS

ON FUNDAMENTAL SYSTEMS OF SYMMETRIC FUNCTIONS†

H. T. ENGSTROM

A set \( S \) of \( n \) polynomials over a field \( K \), symmetric in \( n \) variables, \( x_1, x_2, \ldots, x_n \), is said to form a fundamental system if any rational function over \( K \), symmetric in these variables, can be expressed rationally in terms of the polynomials of \( S \). In this paper we show that any \( n \) algebraically independent symmetric polynomials over a field \( K \) of characteristic zero form a fundamental system if the product of their degrees is less than \( 2n! \).

The result follows from a theorem due to Perron:‡

**Theorem 1.** Between \( n+1 \) polynomials (not constant), \( f_1, f_2, \ldots, f_{n+1} \), in \( n \) variables, of degrees \( m_1, m_2, \ldots, m_{n+1} \), respectively, there is always an identity of the form

\[ \sum C_{r_1r_2 \cdots r_n+1} f_1^{r_1} f_2^{r_2} \cdots f_{n+1}^{r_{n+1}} = 0, \]

where in each term,

\[ \sum_{i=1}^{n+1} m_i \nu_i \leq \prod_{i=1}^{n+1} m_i. \]

* Let every distributive functional on \( S \) be linear, where \( S \) is a topological vector space with the property (Q). If \( S \) is infinite dimensional, let \( \{ x_n \}, (n = 1, 2, \ldots) \), be an infinite set of linearly independent elements. Since \( \lim_{k \to \infty} k^{-1} x_n = \Theta \), we can choose \( y_n \in S, (n = 1, 2, \ldots) \), linearly independent, with \( y_n \to \Theta \). We set \( f(y_n) = 1, f(x) = 0 \) when \( x \) is not a finite linear combination of the \( y_n \), \( f(ax + by) = af(x) + bf(y) \) for any \( x \in S, y \in S \); then \( f \) is a distributive functional on \( S \), and hence is linear on \( S \). Since \( y_n \to \Theta, f(y_n) \to 0 \) as \( n \to \infty \); but this contradicts \( f(y_n) = 1 \). Consequently \( S \) is finite dimensional.

† Presented to the Society, February 25, 1939, under the title A note on fundamental systems of symmetric functions.

‡ O. Perron, Bemerkung zur Algebra, Sitzungsberichte der Bayerischen Akademie, mathematisch-naturwissenschaftliche Abteilung, 1924, pp. 87-101.