JOURNAL OF APPROXIMATION THEORY 23, 335–353 (1978)

Inequalities for Moments and Means

P. C. ROSENBLOOM

Teachers College, Columbia University

Communicated by Oved Shisha

Received January 31, 1977

1. INTRODUCTION

There is an extensive literature on moments of mass distributions

$$f(x) = \int \lambda^x \, d\mu(\lambda), \tag{1}$$

and the corresponding means

$$\begin{aligned} m(x) &= (f(x)/f(0))^{1/x}, \\ m(0) &= \exp(f'(0)/f(0)). \end{aligned} \tag{2}$$

See, for example, Shohat and Tamarkin [13], Rosenbloom [12], Wald [15], Karlin [7], Cargo [2].

If μ is a mass distribution on $R^+ = \{\lambda \mid \lambda \geqslant 0\}$, then its moment function is a Mellin–Stieltjes transform, which is equivalent to a two-sided Laplace–Stieltjes transform (see Widder [16]). It is well known that the mass distribution μ is uniquely determined by f, and there are several inversion formulas. Indeed, under various conditions, μ is already determined by the values of f at the integers, or other denumerable sets (Müntz [10], Szasz [14], Feller [5]), and therefore f is completely determined by its values on such a set. Examples of the nonuniqueness of μ, given the values of f at the nonnegative integers (Pólya and Szegö [11], vol. I, page 134, problem 153) show that some hypothesis is required to ensure the uniqueness.

By Andre Bloch's principle (see [1]), every theorem in the infinite must be a limiting case of theorems in the finite. This suggests the search for inequalities relating such quantities as $\mu(J)$, J any given subinterval of R^+, or $f(x)$, for a given x, to the values of f at any given finite set of points.

One general method of obtaining such inequalities in terms of the values of f at the nonnegative integers is due to Chebyshev, Markov, and Stieltjes, and is presented in [13]. Another approach, applicable to more general interpolation conditions, was given by us in [12].

0021-9045/78/0234–0335$02.00/0

In this paper, we shall work out a few of the simpler applications of our method, and show how a number of recent results (Cargo and Shisha [3, 4], Mond and Shisha [9a, 9b]) follow easily from ours. We also give a few general properties of the extremal moment functions, which may be helpful in treating cases where it is difficult to find explicit formulas for these functions.

Among the applications, we may mention, in particular, that if A is a nonnegative self-adjoint operator on a Hilbert space $\mathscr{H}$, and ξ is any vector in $\mathscr{H}$, then

$$f(x) = (\xi, A^x\xi) = \int \lambda^x \, d\mu(\lambda),$$

where

$$\mu(\lambda) = (\xi, E(\lambda)\, \xi),$$

and E is the spectral measure associated with A. Hence the inequalities obtained here can be applied to analyze the action of A on the cyclic subspace of $\mathscr{H}$ generated by ξ.

2. The Extremal Moment Functons: Summary of Results

We begin by sketching the form which the main results of [12] take, when applied to integrals of the form (1). We shall consider only measures μ with support contained in a given interval $[m, M]$, where $0 \leqslant m < M < +\infty$.

The kernel $K(\lambda, x) = \lambda^x$ is *Cartesian*, that is, if $x_1 < x_2 < \cdots < x_n$, then the number of zeros of the function

$$a_1K(\lambda, x_1) + \cdots + a_nK(\lambda, x_n)$$

in R^+ is at most equal to the number of variations of signs in the sequence $(a_1, ..., a_n)$ of its coefficients. This yields, according to the theory in [12]

Theorem 1. *Let $x_1 < \cdots < x_n$ and $c_1, ..., c_n$ be given, and let $\mathscr{F} = \mathscr{F}(x_1, ..., x_n ; c_1, ..., c_n)$ be the class of moment functions of the form* (1) *satisfying the interpolation conditions*

$$f(x_j) = c_j, \qquad 1 \leqslant j \leqslant n. \tag{3}$$

If $\mathscr{F}$ is nonempty, then it contains two functions φ and ψ such that

$$\varphi(x) \leqslant f(x) \leqslant \psi(x) \tag{4}$$

on the intervals $(x_n, +\infty)$, (x_{n-2}, x_{n-1}), etc., and the reverse inequalities *in the complementary intervals. If, for any x other than the interpolation points x_j, $1 \leqslant j \leqslant n$, $f(x)$ equals $\varphi(x)$ or $\psi(x)$, then that equality holds identically. These extremal moment functions correspond to discrete mass distributions with masses at most n points.*

We can specify the extremal mass distributions more precisely. Let $w(x) = 2$ for $m < x < M$ and $w(m) = w(M) = 1$. We define the *weight* of a set as the sum of the values of $w(x)$ for x in the set. The *weight* of a moment function is defined as the weight of the support of its mass distribution. A moment function of a discrete distribution is an exponential polynomial with positive coefficients. If f is such a moment function, then the weight of f is the maximum of the number of sign-variations in the coefficients of $f - g$, where g ranges over the set of all admissible moment functions of discrete distributions.

THEOREM 2. *The mass distributions associated with the exttemal moment functions in Theorem* 1 *are contained in sets of at most weight n.*

For example, if $n = 1$, so that $\mathscr{F}$ is defined by the single interpolation condition

$$f(x_1) = c_1, \tag{5}$$

then both φ and ψ correspond to masses placed at one endpoint. They must therefore have the forms am^x or AM^x, where a and A are constants. Hence for all f satisfying (5), we have

$$\varphi(x) = f(x_1)\, m^{x-x_1} \leqslant f(x) \leqslant f(x_1)\, M^{x-x_1}$$

for $x > x_1$, and the reverse inequalities for $x < x_1$. This yields the trivial, but useful, results that

$$f(x)/m^x \text{ is increasing or constant,}$$

and

$$f(x)/M^x \text{ is decreasing or constant.}$$

Already for $n = 2$ we obtain nontrivial results. Now one extremal distribution has masses at the two endpoints of the interval $[m, M]$, while the other has a mass at one interior point. Since $\varphi(x) < \psi(x)$ for large x, we must have

$$\psi(x) = Am^x + BM^x$$

and

$$\varphi(x) = ak^x,$$

where A, B, and a are constants, and $m \leqslant k \leqslant M$. An easy computation yields

$$\varphi(x) = f(x_1)(f(x_2)/f(x_1))^{(x-x_1)/(x_2-x_1)} \tag{6}$$

and

$$\psi(x) = \frac{(f(x_1)\, M^{x_2-x_1} - f(x_2))\, m^{x-x_1} + (f(x_2) - f(x_1)\, m^{x_2-x_1})\, M^{x-x_1}}{M^{x_2-x_1} - m^{x_2-x_1}}. \tag{7}$$

The inequality

$$f(x) \leqslant \varphi(x) \qquad \text{for} \quad x_1 \leqslant x \leqslant x_2$$

is equivalent to the well known fact $\log f$ is a convex function. The inequalities

$$\psi(x) \leqslant f(x) \qquad \text{for} \quad x_1 \leqslant x \leqslant x_2$$

and the reverse for $x < x_1$ or $x > x_2$ may be new. The inequalities of Cargo and Shisha [3] and Mond and Shisha [9] can be obtained from this by elementary calculus, and these, in turn, contain well known results of Pólya and Szegö [11, vol. I, pp. 57, 213–214], and Kantorovich [6].

The case $n = 3$ is already more complicated. Now φ and ψ have the forms

$$\begin{aligned} \varphi(x) &= am^x + bk^x, \\ \psi(x) &= AK^x + BM^x, \end{aligned} \tag{8}$$

where the coefficients are nonnegative constants, and k and K are in the interval $[m, M]$. The points k and K, at which the masses are located, are determined by equations which are transcendental, unless $\rho = (x_3 - x_1)/(x_2 - x_1)$ is rational. If $\rho = v/\alpha$, where v and α are relatively prime integers, then we obtain for each of k and K an algebraic equation of degree v, one of whose roots is trivially known. Hence in such a case the determination of k and K can be reduced to the solution of an algebraic equation of degree $v - 1$. For small v we can obtain simple explicit formulas for φ and ψ.

For example, if $\rho = 2$, then we can normalize the various parameters so that $x_1 = -1$, $x_2 = 0$, $x_3 = 1$, $c_2 = 1 = m$. We find that if the support of μ is contained in the half-line $[1, +\infty)$ then

$$f(x) - 1 \geqslant \frac{(1 - f(-1))(f(1) - 1)}{f(1) + f(-1) - 2} \left(\frac{f(1) - 1}{1 - f(-1)}\right)^x$$

for $x > 1$ or $-1 < x < 0$, and the reverse inequality in the complementary intervals. In particular, we obtain the inequality

$$M = \lim_{x \to +\infty} f(x)^{1/x} \geqslant \frac{f(1) - 1}{1 - f(-1)},$$

which gives a lower bound for the range of the distribution. We note that the logarithmic convexity of f only yields the estimate $M \geqslant f(1)$, which is weaker except when all the mass is at one point.

For $n > 3$ the extremal moment functions are still more difficult to determine, except when the x_j are in arithmetical progression.

3. The Cases $n = 1$ and $n = 2$

As mentioned before, the case $n = 1$ reduces to the trivial result:

THEOREM 3. *The ratio $f(x)/m^x$ is increasing or constant, while the ratio $f(x)/M^x$ is decreasing or constant.*

The case $n = 2$ gives the formulas (6) and (7) for the extremal functions φ and ψ. The fact that $f(x)$ is always between $\varphi(x)$ and $\psi(x)$ may also be expressed as follows.

THEOREM 4. *The function* $\log f(x)$ *is a convex function. The mean $m(x)$ is increasing or constant.*

The function

$$r(x) = \frac{f(x)\, M^{-x} - f(0)(m/M)^x}{1 - (m/M)^x} \tag{9}$$

is decreasing or constant. It is constant if and only if all the mass μ is at the endpoints of the interval $[m, M]$.

Since

$$f(x)\, M^{-x} = f(0) - (1 - (m/M)^x)(f(0) - r(x))$$

and $1 - (m/M)^x$ is increasing, Theorem 4 is stronger than the second part of Theorem 3.

Cargo and Shisha [3] solved the problem of finding the maximum of the ratio $m(y)/m(x)$ for given $x, y, x < y$, and given m and M. We now show how their result follows by elementary calculus from Theorem 4.

Without loss of generality, we may assume that $0 < x < y$, $x < y$, and take $x_1 = 0$, $x_2 = x$, $x_3 = y$, and $c_1 = f(0) = 1$. We also set $\rho = y/x$. Then we have

$$\begin{aligned} m(y)^y/m(x)^y &= f(y)/f(x)^\rho \\ &\leqslant \psi(y)/f(x)^\rho, \end{aligned}$$

so that the Cargo–Shisha problem reduces to that of finding the maximum of $\psi(y)/f(x)^\rho$, given x, y, M, and m. We set $f(x) = t$,

$$R = (M^x - m^x)\,\psi(y)/f(x)^\rho = (At - B)\,t^{-\rho}$$

where

$$A = M^y - m^y, \qquad B = (mM)^x\,(M^{y-x} - m^{y-x}).$$

By Theorem 3, t is restricted to the interval $[m^x, M^x]$.

The maximum of R for $t > 0$ is attained at $t = \tau$, where

$$\tau = \rho B/((\rho - 1)\,A).$$

If we set $\gamma = M/m$, then we see that

$$\tau = \frac{\rho m^x \gamma^x(\gamma^{y-x} - 1)}{(\rho - 1)(\gamma^y - 1)}. \tag{10}$$

The fact that τ is an admissible value for t follows from

$$\int_1^\gamma s^{y-x-1}\,ds \leqslant \int_1^\gamma s^{y-1}\,ds \leqslant \gamma^x \int_1^\gamma s^{y-x-1}\,ds.$$

An elementary computation now yields the maximum of R.

COROLLARY 4a (Cargo and Shisha [3]). *If $f(0) = 1$, $0 < x < y$, then*

$$f(y)/f(x)^\rho \leqslant \frac{(\rho - 1)^{\rho-1}\,(\gamma^y - 1)^\rho}{\rho^\rho(\gamma^y - \gamma^x)^{\rho-1}\,(\gamma^x - 1)},$$

where $\rho = y/x$ and $\gamma = M/x$. The equality holds if and only if all the mass μ is at the endpoints of $[m, M]$ and $f(x) = \tau$ in (10).

The other cases considered by Cargo and Shisha ($x < 0 < y$ and $x < y < 0$) can be handled by noticing that

$$f(x) = f(x + x_1)/f(x_1) = \int \lambda^x\,d\mu_1(\lambda),$$

where

$$d\mu_1(\lambda) = f(x_1)^{-1}\,\lambda^{x_1}\,d\mu(\lambda),$$

also is the moment function of a measure with support in $[m, M]$. The limiting cases where x or y is 0 can be handled by using the formula for $m(0)$ in (2).

We can, in a similar way, reduce the problem solved by Mond and Shisha [9] to a problem in elementary calculus. The problem is to find the maximum of the difference $m(y) - m(x)$ for given x, y, $x < y$, and m and M.

We may take the same normalization as before, and set

$$f(x) = (mt)^x.$$

Then we have

$$m(y) - m(x) \leqslant \psi(y)^{1/y} - mt = mh(t),$$

where ψ is given by (7). We find (see Appendix) that $h(1) = h(\gamma) = 0$, $h(t) > 0$ in $(1, \gamma)$, and $h(t) < 0$ outside $[1, \gamma]$. This function h has exactly two relative extrema, a relative minimum at a point $t_1 < 1$ and its absolute maximum, attained at a point t_2 in $(1, \gamma)$.

COROLLARY 4b (Mond and Shisha). *If $f(0) = 1$, $x < y$, $xy \neq 0$, then*

$$f(y)^{1/y} - f(x)^{1/x} \leqslant mh(t_2), \tag{11}$$

where

$$\begin{aligned} h(t) &= (A + Bt^x)^{1/y} - t, \\ A &= (\gamma^x - \gamma^y)/(\gamma^x - 1), \\ B &= (\gamma^y - 1)/(\gamma^x - 1), \end{aligned} \tag{12}$$

and t_2 is the unique point in the interval $(1, \gamma)$ where h attains its maximum. If $y \neq 1$, then t_2 is the unique zero of

$$J(t) = A + Bt^x - C_1 t^{(1-x)y/(1-y)},$$

$$C_1 = (y/(xB))^{y/(1-y)},$$

in the interval $[1, \gamma]$. For $y < 1$, J has one other positive zero, and this is in the interval $(0, 1)$.

The equality in (11) is attained only when μ is the discrete distribution with mass

$$w = (t_2{}^x - 1)/(\gamma^x - 1) \tag{13}$$

at M and mass $1 - w$ at m.

If $y = 1$, then t_2 is the positive root of

$$xBt^{x-1} = 1.$$

Clearly both corollaries are special cases of

COROLLARY 4c. *Let $G(X, Y)$ be monotonically increasing in Y for $0 < X < Y$. Then if $f(0) = 1$, $x < y$, we have*

$$\max G(f(x)^{1/x}, f(y)^{1/y}) = \max_{1 \leqslant v \leqslant y} G(mv, m(A + Bv^x)^{1/y}),$$

where A and B are given in (12). *The maximum is attained only when μ is a discrete distribution with masses w at M and $1 - w$ at m, where w is given by* (13), *and v is a point where $G(mv, m(A + Bv^x)^{1/y})$ attains its maximum in* $[1, \gamma]$.

We now wish to interpret these inequalities. In the following, the quantity

$$\alpha = \mu(\{M\}) = \alpha(f),$$

the mass placed at the point M, plays an important role. We begin with an elementary remark, which must be well known.

THEOREM 5. *If f is given by* (1), *then* $\lim_{x\to+\infty} f(x)\, M^{-x} = \alpha$.

Proof. Obviously we have

$$f(x) \geqslant \alpha M^x.$$

On the other hand, we see that for $m < a < M$,

$$f(x) \leqslant \mu([a, M])\, M^x + \mu([m, a))\, a^x,$$

so that

$$\alpha \leqslant \lim_{x\to+\infty} f(x)\, M^{-x} \leqslant \mu([a, M]).$$

We now let a approach M^-, and obtain the stated result.

Thus the ratio $f(x)\, M^{-x}$ decreases to α as x increases, unless all the mass is at M and the ratio is constant. Theorem 4 says that the ratio $r(x)$, in (9), also decreases to α, unless all the mass is at the two points m and M and then this ratio is constant. This gives a sharper upper bound for α, using the additional information of the values of m and $f(0)$, as well as that of $f(x)$. For we have

$$r(x) < f(x)\, M^{-x},$$

unless all the mass is at M.

For further reference we note the following alternative form of the inequality $f(y) \leqslant \psi(y)$ in the present case of $n = 2$:

if $x_1 < x_2 < x_3$, then

$$0 \leqslant \begin{vmatrix} m^{x_1} & f(x_1) & M^{x_1} \\ m^{x_2} & f(x_2) & M^{x_2} \\ m^{x_3} & f(x_3) & M^{x_3} \end{vmatrix}.$$

We do not know any really good analog for $n > 2$.

4. The Case $n = 3$

In order to avoid a profusion of subscripts, let us set $x_1 = 0$, $x_2 = x$, $x_3 = y$. Without loss of generality, we may assume $c_1 = f(0) = 1$. Now φ and ψ are given by (8), where k and K are determined by the equations

$$\begin{vmatrix} 1 & 1 & 1 \\ m^x & k^x & f(x) \\ m^y & k^y & f(y) \end{vmatrix} = 0$$

and (14)

$$\begin{vmatrix} 1 & 1 & 1 \\ K^x & M^x & f(x) \\ K^y & M^y & f(y) \end{vmatrix} = 0,$$

respectively. These equations can be put in the alternative forms

$$\frac{f(y)\, m^{-y} - 1}{(k/m)^y - 1} = \frac{f(x)\, m^{-x} - 1}{(k/m)^x - 1}$$

and (15)

$$\frac{1 - f(y)\, M^{-y}}{1 - (K/M)^y} = \frac{1 - f(x)\, M^{-x}}{1 - (K/M)^x},$$

or

$$\frac{(k/m)^t - 1}{(k/m)^x - 1} = \frac{f(y)\, m^{-y} - 1}{f(x)\, m^{-x} - 1},$$

and (16)

$$\frac{1 - (K/M)^y}{1 - (K/M)^x} = \frac{1 - f(y)\, M^{-y}}{1 - f(x)\, M^{-x}}.$$

If we set $\rho = y/x > 1$ and

$$\begin{aligned} g(\lambda) &= \frac{\lambda^\rho - 1}{\lambda - 1} \\ &= \rho \int_0^1 (1 + s(\lambda - 1))^{\rho-1}\, ds, \end{aligned} \tag{17}$$

we see that g is strictly increasing for $\lambda \geqslant 0$. Since, by Theorem 4, the ratio

$$r = \frac{f(y)\, m^{-y} - 1}{f(x)\, m^{-x} - 1}$$

is greater than $g(1) = \rho$ unless all the mass is at m, we find that the equation $g(\lambda) = r$ has a unique positive root λ_0, and $\lambda_0 > 1$. Then k in (16) is uniquely determined as

$$k = m\lambda_0^{1/x}.$$

Theorem 4 also implies that

$$r < g((M/m)^x),$$

so that

$$k < M,$$

unless all the mass is at m and M.

Similarly we show that K is uniquely determined and is in the interval $[m, M]$. It is an endpoint of the interval if and only if all of the mass μ is at the endpoints.

We note also that $\psi - \varphi$ has zeros at 0, x, and y, so that there must be at least three sign-variations in the coefficients of this exponential polynomial, unless ψ and φ are identical. Hence we find that

$$K < k,$$

unless all the mass is at one point.

THEOREM 6. *If $f(0) = 1$, then the inequalities*

$$\varphi(z) \leqslant f(z) \leqslant \psi(z) \tag{18$_n$}$$

hold in $(0, x)$ and $(y, +\infty)$, and the reverse inequalities in $(-\infty, 0)$ and (x, y), where φ and ψ have the form (8), and k and K are the unique positive roots of (16). If any of the equalities in $(18)_n$ holds at any point other than 0, x, and y, then that equality holds identically.

The positive roots of (16) satisfy

$$m < K < k < M,$$

unless all the mass is at the endpoints of $[m, M]$.

The inequality $f(z) < \psi(z)$ can be put in the form:

COROLLARY 6a. *If K is determined by (16), then*

$$s(z) = \frac{f(z)\, M^{-z} - (K/M)^z}{1 - (K/M)^z} \leqslant s(y) \tag{18}$$

for $z > y$ or $0 < z < x$. The mass α at M satisfies

$$\alpha \leqslant s(y).$$

Actually this result can be sharpened.

COROLLARY 6b. *The ratio* $s(z)$ *in* (18) *is decreasing or constant for* $z > y$ *and increasing or constant for* $z < x$. *It is constant if and only if all the mass* μ *is at the points* K *and* M.

Proof. We may assume that s is not constant, so that not all the mass is at the points K and M. If there are z_1 and u such that $y < z_1 < u$ and $s(u) \geqslant s(z_1)$, then there is a z in $(y, z_1]$ such that $s(z) = s(u)$. Then K is also the solution of (16) with x and y replaced by z and u, respectively. Let ψ_1 be the extremal function for the interpolation with prescribed values at u, z and u. It follows from Theorem 6 that

$$f(t) < \psi_1(t) \qquad \text{for} \quad 0 < t < z,$$

and for $t = y$ this contradicts (18).

We can treat the intervals $(0, x)$ and $(-\infty, 0)$ in the same way.

Thus the ratio $s(z)$ decreases to α as z increases to $+\infty$. Since $m < K$ unless all the mass is at two points, we see that

$$s(z) < r(z) \qquad \text{for} \quad z > y$$

so that $s(y)$ gives a sharper estimate for α than is given in Theorem 4.

It is of interest that ψ is independent of m, and that K provides an upper estimate of m. If g is defined by (17), then the inequality

$$f(x)^{1/x} \leqslant f(y)^{1/y}$$

implies that

$$g(f(x)/M^x) \geqslant g((K/M)^x).$$

Hence we have

$$K \leqslant f(x)^{1/x},$$

and the equality holds if and only if all the mass is at one point. Thus the upper estimate K for m, in terms of the data $f(x)$, $f(y)$, and M, is sharper than the estimates using $f(x)$ alone.

Similarly we can prove

COROLLARY 6c. *The ratio*

$$s_1(z) = \frac{f(z)\, m^{-z} - 1}{(k/m)^z - 1}$$

is increasing or constant for $z > y$ *and decreasing or constant for* $z < x$. *It is constant if and only if all the mass is at the points* m *and* k.

The special case $z = 0$ may be of interest:

COROLLARY 6d. *The inequality*

$$s_1(z) \leqslant s_1(0) = \frac{f'(0) - \log m}{\log(k/m)}$$

holds for $0 < z < x$ *and the reverse for* $z < 0$.

From the monotonicity of g in (17) and the inequality

$$g(1 - s^{-1}) < s < g(s^{1/(\rho-1)}) \qquad \text{for} \quad s > 1,$$

we obtain these estimates

$$K \geqslant M \left(\frac{f(x)\, M^{-x} - f(y)\, M^{-y}}{1 - f(y)\, M^{-y}}\right)^{1/x}, \tag{19}$$

and

$$k \leqslant m \left(\frac{f(y)\, m^{-y} - 1}{f(x)\, m^{-x} - 1}\right)^{1/(y-x)}$$

Both are asymptotic equalities for large y, if m, M and x are held constant.

For $\rho = 2$, 3, and 3/2 we can give simple explicit formulas for k and K in terms of the quantities

$$S = \frac{1 - f(y)\, M^{-y}}{1 - f(x)\, M^{-x}},$$

and

$$s = \frac{f(y)\, m^{-y} - 1}{f(x)\, m^{-x} - 1}.$$

We obtain:

for $\rho = 2$,

$$K = M(S - 1)^{1/x}, \quad k = m(s - 1)^{1/x};$$

for $\rho = 3$,

$$K = M\sigma^{1/x}, \quad k = m\tau^{1/x},$$
$$g(\sigma) = \sigma^2 + \sigma + 1 = S;$$
$$g(\tau) = s, \quad 0 < \sigma, \tau;$$

and for $\rho = 3/2$,

$$K = K\sigma^{2/x}, \quad k = m\tau^{2/x},$$
$$g(\sigma^2) = (\sigma^2 + \sigma + 1)/(\sigma + 1) = S,$$
$$g(\tau^2) = s.$$

5. The Case $n = 4$

We may take the interpolation points as 0, x, y, z, in increasing order. Now the extremal functions have the forms

$$\begin{aligned} \psi(t) &= Am^t + BR^t + CM^t, \\ \varphi(t) &= ar_1{}^t + br_2{}^t. \end{aligned} \tag{20}$$

The constants R, r_1, and r_2, are determined by the following equations:

$$h(R) = \begin{vmatrix} 1 & 1 & 1 & 1 \\ m^x & R^x & M^x & f(x) \\ m^y & R^y & M^y & f(y) \\ m^z & R^z & M^z & f(z) \end{vmatrix} = 0, \tag{21}$$

$$\begin{vmatrix} 1 & 1 & 1 \\ r_1{}^x & r_2{}^x & f(x) \\ r_1{}^y & r_2{}^y & f(y) \end{vmatrix} = \begin{vmatrix} 1 & 1 & 1 \\ r_1{}^x & r_2{}^x & f(x) \\ r_1{}^z & r_2{}^z & f(z) \end{vmatrix} = 0. \tag{22}$$

These are transcendental equations unless x, y, and z are commensurable. The function $h(\exp t)$ is an exponential polynomial whose sequence of coefficients has 3 sign-variations. Hence it has one or three real roots. Since $h(m) = h(M) = 0$, therefore h has just one other positive root, which must be between m and M.

The system (22) has at most one solution with $0 < r_1 < r_2$. For if there were two, there would be two functions of the form of φ in (20) satisfying the interpolation conditions at 0, x, y, and z. But their difference would have at most 3 sign-variations in the coefficients, and therefore less than 4 non-negative roots. Since an extremal function φ exists, then there is a unique positive solution (r_1, r_2) of (22), and $m \leqslant r_1 < r_2 \leqslant M$. (If $r_1 = r_2$, then this moment function of a mass at one point is the only one satisfying the interpolation conditions, and $\varphi = \psi$.)

Since $\psi - \varphi$ has at least 4 nonnegative roots, then there must be at least four sign-variations in the coefficients. We infer that

$$m < r_1 < R < r_2 < M,$$

unless all the mass is at two points, one of which is an endpoint.

We note that (22) is independent of m and M, so that φ is the extremal function which minimizes $f(t)$, for given $t > z$, among *all* moment functions satisfying the interpolation conditions. It also minimizes $f(t)$ for t in the intervals $(-\infty, 0)$ and (x, y), and maximizes $f(t)$ in the complementary intervals $(0, x)$ and (y, z). The condition $m \leqslant r_1 < r_2 \leqslant M$ is necessary and sufficient that there exist any moment function satisfying the interpolation conditions, whose mass has support in $[m, M]$.

An equivalent condition which may be useful when r_1 and r_2 are difficult to determine can be given. Let φ_1 and ψ_1 be the extremal functions discussed in Section 3 corresponding to the interpolation conditions at 0, x, and y, and let k and K be the constants defined there (Eqs. (14), (15)). Then for there to be a moment function, whose mass has support in $[m, M]$, satisfying also the interpolation condition at z, it is necessary and sufficient that the prescribed value for $f(z)$ satisfy

$$\varphi_1(z) \leqslant f(z) \leqslant \psi_1(z).$$

Examples in which k and K are given by simple explicit formulas are given in the previous section.

By simple considerations on the sign-variations of the coefficients, we can easily establish the following order relations:

$$r_1 \leqslant K \leqslant R \leqslant k \leqslant r_2 .$$

If any equality sign holds here, then all of the mass μ is at one or two points.

The simplest case in which we can give explicit formulas for R, r_1, and r_2 is that of $y = 2x$, $z = 3x$. In this case, we have

$$mMR = \left\{ \begin{vmatrix} m^x & f(x) & M^x \\ m^{2x} & f(2x) & M^{2x} \\ m^{3x} & f(3x) & M^{3x} \end{vmatrix} \Bigg/ \begin{vmatrix} 1 & 1 & 1 \\ m^x & f(x) & M^x \\ m^{2x} & f(2x) & M^{2x} \end{vmatrix} \right\}^{1/x}$$

and $r_1{}^x$ and $r_2{}^x$ are the solutions of the quadratic equation

$$Q(t) = t^2 - ut + v = 0,$$

$$u = \frac{f(3x) - f(x)\,f(2x)}{f(2x) - f(x)^2},$$

$$v = \frac{f(x)\,f(3x) - f(2x)^2}{f(2x) - f(x)^2}.$$

The inequality $m^x \leqslant r_1{}^x \leqslant r_2{}^x \leqslant M^x$ implies that

$$(r_2{}^x - r_1{}^x)^2 \leqslant (M^x - m^x)^2,$$

or

$$u^2 - 4v \leqslant 4(M^x - m^x)^2.$$

This lower estimate for the range, as well as the relation

$$4v \leqslant u^2$$

or

$$4(f(2x) - f(x)^2)(f(x)\,f(3x) - f(2x)^2) \leqslant (f(3x) - f(x)\,f(2x))^2,$$

may be new. The equality holds here if and only if all the mass μ is at one point.

The value of C in (20) gives the estimate

$$\alpha \leqslant \begin{vmatrix} 1 & 1 & 1 \\ m^x & R^x & f(x) \\ m^y & R^y & f(y) \end{vmatrix} \Bigg/ \begin{vmatrix} 1 & 1 & 1 \\ m^x & R^x & M^x \\ m^y & R^y & M^y \end{vmatrix} \tag{23}$$

for the mass at the point M.

We sum up the main results of this chapter.

THEOREM 7. *Equation* (21) *has a unique positive root R other than m and M, and the system* (22) *has a unique positive solution (r_1, r_2) with $r_1 < r_2$, unless all the mass is at one point. Unless all the mass is at one or two points, we have the order relation*

$$m < r_1 < K < R < k < r_2 < M,$$

where k and K are the solutions of (14). *If φ and ψ are defined by* (20), *then*

$$\varphi(t) \leqslant f(t) \leqslant \psi(t)$$

in the intervals $(-\infty, 0)$, (x, y), $(z, +\infty)$, and the reverse inequality holds in the complementary intervals. If an equality holds for any t other than 0, x, y, or z, then that equality holds identically and all the mass is at two or three points. The mass α at the point M satisfies (23).

6. THE STRUCTURE OF THE EXTREMAL FUNCTIONS

We now wish to establish some general results concerning the general structure of the extremal functions, which sharpen Theorems 1 and 2. The properties discussed here are illustrated by the phenomena occurring in the special cases treated in the preceding sections. We use the notation of Theorem 1.

THEOREM 8. *If $\varphi \not\equiv \psi$, then the coefficient of M^y in $\psi(y)$ is positive*:

$$\alpha(\psi) = \lim_{y\to+\infty} \psi(y)\, M^{-y} > 0,$$

so that ψ is the moment function of a distribution with positive mass at M.

Proof. We proceed by induction. The assertion is trivially true for $n = 1$. Let $\mathscr{F}_{n-1} = \mathscr{F}(x_1, ..., x_{n-1}; c_1, ..., c_{n-1})$ be the set of moment functions satisfying the interpolation conditions at x_j, $j < n$, and let φ_{n-1}

and ψ_{n-1} be the extremal functions of $\mathscr{F}_{n-1}$. Since $\mathscr{F}_{n-1} \subset \mathscr{F}$, we see that

$$\varphi_{n-1}(y) \leqslant \varphi(y) < \psi(y) \leqslant \psi_{n-1}(y)$$

for $y > x_n$.

By the induction hypothesis, we know that $\alpha(\psi_{n-1})$ is positive. Furthermore, we must have $\varphi_{n-1}(y) < \psi_{n-1}(y)$ for $y > x_{n-1}$ and, in particular, $\varphi_{n-1}(y) < \psi_{n-1}(y)$.

Now if $c_n = \varphi(x_n) = \psi(x_n)$ is equal to $\varphi_{n-1}(x_n)$ or $\psi_{n-1}(x_n)$, then φ and ψ must coincide with that extremal function of $\mathscr{F}_{n-1}$, by Theorem 1. It follows that

$$\varphi_{n-1}(x_n) < c_n < \psi_{n-1}(x_n).$$

We set

$$\lambda = (c_m - \varphi_{n-1}(x_n))/(\psi_{n-1}(x_n) - \varphi_{n-1}(x_n))$$

and

$$f = (1 - \lambda)\, \varphi_{n-1} + \lambda\psi_{n-1}. \tag{24}$$

Then f is in $\mathscr{F}$, so that $f(y) \leqslant \psi(y)$ for $y > x_n$, and $\alpha(\psi) \geqslant \alpha(f) \geqslant \lambda\alpha(\psi_{n-1}) > 0$.

Remark. Using the notation of this proof, we observe that if $\varphi \equiv \psi$ but $\varphi_{n-1} \not\equiv \psi_{n-1}$, and if $\alpha(\psi) = 0$, then $\psi \equiv \varphi_{n-1}$. For the f defined by (24) is in $\mathscr{F}$, and so must coincide with ψ. Hence λ must be 0, i.e., $c_n = \psi(x_n) = \varphi_{n-1}(x_n)$, and then ψ coincides with φ_{n-1}.

COROLLARY 8a. *If $\varphi \not\equiv \psi$, then*

$$\lim_{y\to-\infty} \psi(y)\, m^{-y} > 0 \qquad \textit{for} \quad n \textit{ even},$$

$$\lim_{y\to-\infty} \varphi(y)\, m^{-y} > 0 \qquad \textit{for} \quad n \textit{ odd}.$$

Proof. If f is defined by (1), then

$$f_1(x) = f(-x) = \int \lambda^x \, d\mu_1(\lambda),$$

where $d\mu_1(\lambda) = d\mu(\lambda^{-1})$. Thus f_1 is the moment function of a mass distribution confined to the interval $[M^{-1}, m^{-1}]$, and belongs to the set $\mathscr{F}(-x_{n,m}, -x_1; c_{n,m}, c_2)$. Also the extremal functions of this set are $\varphi(-x)$ and $\psi(-x)$ if n is even, and the reverse if n is odd. The corollary now follows from the theorem.

COROLLARY 8b. *The points where ψ and φ have positive masses mutually separate each other, unless $\psi \equiv \varphi$, and then $\mathscr{F}$ has only one member.*

Proof. We may assume that $\psi \not\equiv \varphi$. Suppose that n is even. Then ψ has positive masses at m and M and at no more than $(n-2)/2$ interior points of the interval $[m, M]$. If all the masses of φ are at interior points, then there can be at most $n/2$ of these. Thus the difference $\psi - \varphi$ contains at most

$$2 + (n-2)/2 + (n/2) = n + 1$$

terms, and so can have at most n sign-variations in its coefficients. But it has at least n real roots and therefore at least this many sign-variations. Therefore there are exactly $n + 1$ terms, and the coefficients have alternating signs.

If φ has a mass at either endpoint, then that mass is not more than that of ψ at the same point. Thus the difference $\psi - \varphi$ has a nonnegative mass at such an endpoint. Since φ can have at most $(n-2)/2$ masses at interior points, the difference cannot have more than n terms, which is impossible.

The case of odd n is treated similarly. Now each of the extremal functions has positive masses at one endpoint and $(n-1)/2$ interior points.

The discussion, in the course of this proof, of the number of terms in $\psi - \varphi$ yields also

COROLLARY 8c. *If $\psi \not\equiv \varphi$, then both have exactly weight n.*

Remark. We can now improve the previous remark. If $\varphi \equiv \psi$ but $\varphi_{n-1} \not\equiv \psi_{n-1}$, and $\alpha(\psi) > 0$, then $\psi \equiv \psi_{n-1}$. For if not, λ in (24) is positive, and f has positive masses at both endpoints and at $n-2$ interior points. Hence the weight of ψ (which coincides with f) is $2 + 2(n-2) = 2n - 2$, and this is greater than n if $n > 2$. The case $n = 2$ follows from the explicit formulas in (6) and (7).

Let $\mathscr{F}_k = \mathscr{F}(x_1, \ldots, x_k ; c_1, \ldots, c_k)$ and let φ_k and ψ_k be the extremal functions of $\mathscr{F}_k$.

THEOREM 9. *If the weight of ψ is k, $k < n$, then ψ coincides with φ_k or ψ_k.*

Proof. We have $\varphi_k(y) \leqslant \psi(y) \leqslant \psi_k(y)$ for $y > x_k$. If $\psi \not\equiv \psi_k$, then $\psi(y) < \psi_k(y)$ for $y > x_k$. Since the weight of ψ is k, then the number of real roots of $\psi - f$, for any moment function f of a discrete distribution with support in $[m, M]$, is at most k, unless $\psi \equiv f$. Hence for $f \in \mathscr{F}_k$, $f \not\equiv \psi$, this difference has only the k zeros $x_1, \ldots, x_k$. Let

$$\lambda = \frac{c_{k+1} - \varphi_k(x_{k+1})}{\psi_k(x_{k+1}) - \varphi_k(x_{k+1})},$$

and

$$f = (1 - \lambda)\, \varphi_k + \lambda \psi_k .$$

Since $f(x_{k+1}) = \psi(x_{k+1})$ and $f \in \mathscr{F}_k$, it follows that $f \equiv \psi$. If $0 < \lambda < 1$, then f has exactly $k + 1$ terms, corresponding to positive masses at m and M and at $k - 1$ interior points. Thus the weight of f is

$$2 + 2(k - 1) = 2k > k,$$

which gives us a contradiction. Therefore $\psi(x_{k+1}) = c_{k+1}$ must equal $\varphi_k(x_{k+1})$, since $\lambda < 1$, and our assertion follows.

If the interpolation points x_j are in arithmetic progression, then the points, at which either extremal function has positive masses, are the zeros of a quasiorthogonal polynomial. The interlacing property of these points, expressed in Corollary 8b, reduces to the classical separation property of the zeros of these polynomials. (See [13], pp. 36–38.)

Appendix

We give here a brief discussion of the function h defined in (12), which completes in some minor respects, that of Marshall and Olkin [8] which is used by Mond and Shisha [9]. We shall treat here the case where $0 < x < y$. The other cases can be handled in a similar way.

If $0 < x < y$, then $A < 0 < B$ in (12). Then $h(t)$ is defined only for $A + Bt^x \geqslant 0$, that is,

$$t \geqslant t_0 = (-A/B)^{1/x}.$$

This function $h(t)$ is positive if and only if the power sum

$$H(t) = A + Bt^x - t^y > 0.$$

Since there are two sign-variations in the coefficients of H, H can have at most two nonnegative zeros. Since $H(1) = H(\gamma) = 0$, there are no others. Also H is negative for small or large t. Hence H is negative in $(0, 1)$ and in $(\gamma, +\infty)$, and positive in $(1, \gamma)$. Therefore also h is positive in $(1, \gamma)$ and negative outside this interval. By Rolle's theorem h' has a zero in $(1, \gamma)$.

For $y \neq 1$, an elementary computation shows that if $h'(t) = 0$, then

$$J(t) = A + Bt^x - C_1 t^z = 0,$$

where

$$z = y(1 - x)/(1 - y),$$

and C_1 is the positive number defined in the statement of Corollary 4b. If $x < 1 < y$, then $z < 0$ while if $1 < x < y$, then $0 < z < x$, and in either case J, and therefore h' has at most one nonnegative root. If $y < 1$, then

$x < z$ and J has at most two nonnegative zeros. But if $y < 1$, then $h'(t_0) = -1 < 0$, so that h has a relative minimum in $(t_0, 1)$. Then h' has exactly two zeros, one in $(t_0, 1)$ and one in $(1, \gamma)$.

References

1. A. Bloch, Les fonctions holomorphes et méromorphes dans le cercle-unité, *Mémor. Sci. Math.* **20** (1926).
2. G. T. Cargo, An elementary, unified treatment of complementary inequalities, *in* "Inequalities" (O. Shisha, Ed.), Vol. III, pp. 39–64, Academic Press, New York, 1972.
3. G. T. Cargo and O. Shisha, Bounds on ratios of means, *J. Res. Nat. Bur. Standards Sect. B.* **66B** (1962), 169–170.
4. G. T. Cargo and O. Shisha, A metric space connected with generalized means, *in* "Inequalities" (O. Shisha, Ed.), Vol. II, pp. 163-178, Academic Press, New York, 1970.
5. W. Feller, On Müntz' theorem and completely monotone functions, *Amer. Math. Monthly* **75** (1968), 342–350.
6. L. V. Kantorovich, Functional analysis and applied mathematics (in Russian), *Uspehi Mat. Nauk* **3** (1948), 89–185.
7. S. Karlin, "Total Positivity," Stanford Univ. Press, Stanford, Calif. 1968.
8. A. W. Marshall and I. Olkin, Reversal of the Lyapunov, Hölder, and Minkowski inequalities and other extensions of the Kantorovich inequality, *J. Math. Anal. Appl.* **8** (1964), 503–14.
9. B. Mond and O. Shisha. (a). Ratios of means and applications, pp. 191–197, and (b). Bounds on differences of means, pp. 293–308, *in* "Inequalities" (O. Shisha, Ed.). Vol. II Academic Press, New York, 1967.
10. Ch. H. Müntz, Über den Approximationssatz von Weierstrass, *in* "Math. Abhandlungen, H. A. Schwarz zu seinem 50, Doktorjubileum gewidmet, Berlin, (1914)," pp. 303–12.
11. G. Pólya and G. Szegö, "Problems and Theorems in Analysis," Springer, New York/ Berlin, 1972.
12. P. C. Rosenbloom, Quelques classes de problèmes extremaux, *Bull. Soc. Math. France* **79** (1951), 1–58 and **80** (1952), 183–215.
13. J. A. Shohat and J. D. Tamarkin, "The Problem of Moments," Math. Surveys, No. 1, *Amer. Math. Soc.*, New York, 1943.
14. O. Szasz, Über die Approximation steiger Funktionen durch lineare Aggregate von Potenzen, *Math. Ann.* **77** (1916), 482–496.
15. A. Wald, Limits of a distribution function determined by absolute moments and inequalities satisfied by absolute moments. *Trans. Amer. Math. Soc.* **46** (1939), 280–306.
16. D. V. Widder, "The Laplace Transform," Princeton Univ. Press, Princeton, N.J., 1941.