A VARIANT OF CARATHÉODORY'S PROBLEM *

by GILBERT STRANG

(Received 23rd May 1967)

1. In this note we ask two questions and answer one. The questions can be combined as follows:

Does there exist a polynomial of the form

$$p(z) = \Sigma c_j(z-1)^j \tag{1}$$

which starts with prescribed complex coefficients $c_0, \ldots, c_{r-1}$, and satisfies

$$\text{I:} \quad \operatorname{Re} p(z) > 0 \text{ for } |z| \leqq 1,\ z \neq 1?$$

$$\text{II:} \quad |p(z)| < 1 \text{ for } |z| \leqq 1,\ z \neq 1?$$

These differ from the classical problems of Carathéodory in one essential respect: the values of p and its first $r-1$ derivatives are given at the point $z = 1$ on the circumference of the unit circle, while in the original problem they were given at $z = 0$. Carathéodory's own answer was in terms of his " moment curve ", but the forms studied a few years later by Toeplitz yield a more convenient statement of the solution. Since we want to reduce question I to Carathéodory's first problem, we recall the classical result:

There exists a polynomial $P(z) = \Sigma a_j z^j$ starting with prescribed coefficients $a_0, \ldots, a_{q-1}$ and satisfying $\operatorname{Re} P > 0$ for $|z| \leqq 1$ if and only if the associated Toeplitz form is positive definite: whenever $v \neq 0$,

$$(T_{q-1}v, v) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \operatorname{Re} \sum_0^{q-1} a_j e^{ij\theta} \left| \sum_0^{q-1} v_k e^{ik\theta} \right|^2 d\theta > 0. \tag{2}$$

It is easy to see why (2) is necessary. If there is such a polynomial P, then for $v \neq 0$,

$$0 < \frac{1}{2\pi}\int \operatorname{Re} P(e^{i\theta}) \left| \sum_0^{q-1} v_k e^{ik\theta} \right|^2 d\theta = (T_{q-1}v, v);$$

the other terms $a_q e^{iq\theta} + \ldots + a_Q e^{iQ\theta}$ in P contribute nothing to the integral.

In stating the sufficiency of (2) we have taken some liberties with the more delicate result derived by Grenander and Szegö [**1**, p. 151]. They produce a power series $f(z) = \sum_0^{\infty} a_j z^j$ regular with $\operatorname{Re} f \geqq 0$ for $|z| < 1$, whenever T_{q-1} is a non-negative form. To construct our P, suppose T_{q-1} is in fact positive definite. Then it remains so if a_j is replaced by $a_j' = a_j(1+\varepsilon)^j$, $1 \leqq j < q$ and

* This research was supported by the Sloan Foundation, the National Science Foundation (GP 7477), and the Office of Naval Research.

https://doi.org/10.1017/S0013091500012165 Published online by Cambridge University Press

$a_0' = a_0 - \varepsilon$, for a suitably small $\varepsilon > 0$. Now [1] provides a power series $f'(z')$ starting with the a_j' and satisfying $\mathrm{Re} f' \geqq 0$ for $|z'| < 1$. Replacing z' by $z/(1+\varepsilon)$, we have a power series f starting with the a_j, regular in $|z| < 1+\varepsilon$, and satisfying $\mathrm{Re} f \geqq \varepsilon$ in this circle. Truncating the series f at sufficiently large Q gives the polynomial P.

In short, one can decide after a fixed number of computations with the a_j whether or not the required polynomial P exists. It is an answer of this sort, in terms of $c_0, \ldots, c_{r-1}$, that we want for our problems. We have elsewhere investigated several special cases of questions I and II, in connection with difference schemes for mixed initial-boundary value problems [2-4]. Our methods of proof were very much *ad hoc*, however, and a more systematic treatment seems justified.

One could also think of replacing (1) by

$$P(z) = \Sigma c_j (z - z_0)^j$$

for points z_0 other than 1 or 0. In case $|z_0| = 1$ or $|z_0| < 1$, the obvious conformal map of the unit circle onto itself transforms the problem to one of the two problems already described. For $|z_0| > 1$, it is easy to show that the required polynomial always exists.

2. We begin with the calculation on which our solution depends.

Lemma 1. *The space of polynomials* $\sum_r^R c_j(e^{i\theta}-1)^j$ *coincides for* $r = 2s$ *with the space of functions of the form* $(1-\cos\theta)^s \sum_s^{R-s} a_k e^{ik\theta}$.

Proof. Both are (complex) vector spaces of dimension $R-r+1$. To prove that they coincide, we have only to show that the second contains the first. For $r \leqq j \leqq R$ we have

$$\begin{aligned}(e^{i\theta}-1)^j &= (e^{i\theta/2} - e^{-i\theta/2})^r e^{ir\theta/2}(e^{i\theta}-1)^{j-r} \\ &= (1-\cos\theta)^s(-2)^s e^{is\theta}(e^{i\theta}-1)^{j-2s}\end{aligned}$$

and the right side lies in the second vector space. Therefore the same is true for any linear combination of the powers $(e^{i\theta}-1)^j$, $r \leqq j \leqq R$, completing the proof.

If r is even, this result almost reduces our question I to Carathéodory's problem. We are looking for $c_r, \ldots, c_R$ such that

$$\mathrm{Re}\left[\sum_0^{r-1} c_j(e^{i\theta}-1)^j + \sum_r^R c_j(e^{i\theta}-1)^j\right] > 0 \text{ for } \theta \neq 0 \ (\mathrm{mod}\ 2\pi). \tag{3}$$

According to the lemma, this is equivalent to looking for $a_s, \ldots, a_{R-s}$ such that

$$\frac{\mathrm{Re}\sum_0^{r-1} c_j(e^{i\theta}-1)^j}{(1-\cos\theta)^s} + \mathrm{Re}\sum_s^{R-s} a_k e^{ik\theta} > 0 \text{ for } \theta \neq 0 \ (\mathrm{mod}\ 2\pi). \tag{4}$$

Admitting the possibility that a factor $(1-\cos\theta)^t$ might cancel in the first term, we need the following result.

Lemma 2. *Suppose that $f(\theta)$ is a real trigonometric polynomial, $f(0)>0$, and $0 \leqq t<s$. Then there exist finitely many coefficients $a_s, \ldots, a_S$ such that*

$$\frac{f(\theta)}{(1-\cos\theta)^{s-t}}+\mathrm{Re}\sum_s^S a_k e^{ik\theta}>0 \text{ for } \theta \neq 0 \pmod{2\pi} \tag{5}$$

if and only if the Toeplitz form

$$(T_{t-1}(f)u, u)=\frac{1}{2\pi}\int_{-\pi}^{\pi} f(\theta)\left|\sum_0^{t-1} u_k e^{ik\theta}\right|^2 d\theta \tag{6}$$

is positive definite. If $t=0$ this condition is vacuous and (5) *can always be satisfied.*

Proof. Suppose that (6) were not positive definite. Then for some polynomial $P=\Sigma u_k e^{ik\theta}$ of degree less than t (we shall always normalize to $\Sigma\mid u_k\mid^2=1$) we have

$$\int f(\theta)\mid P(\theta)\mid^2 d\theta \leqq 0.$$

For any choice of the a_k, this implies

$$0 \geqq \int \frac{f(\theta)}{(1-\cos\theta)^{s-t}}(1-\cos\theta)^{s-t}\mid P\mid^2 d\theta$$

$$=\int\left[\frac{f(\theta)}{(1-\cos\theta)^{s-t}}+\mathrm{Re}\sum a_k e^{ik\theta}\right](1-\cos\theta)^{s-t}\mid P\mid^2 d\theta \tag{7}$$

since $(1-\cos\theta)^{s-t}\mid P\mid^2$ is of degree $<s$. Clearly (5) cannot hold if (7) does.

For the converse, suppose that the form (6) is positive definite; for all (normalized) u_k,

$$\int \frac{f(\theta)}{(1-\cos\theta)^{s-t}}(1-\cos\theta)^{s-t}\left|\sum_0^{t-1} u_k e^{ik\theta}\right|^2 d\theta>0. \tag{8}$$

We claim that there is a trigonometric polynomial g, such that

$$g(\theta) \leqq f(\theta)/(1-\cos\theta)^{s-t} \text{ for all } \theta,$$

for which the form

$$\int g(\theta)\left|\sum_0^{s-1} v_k e^{ik\theta}\right|^2 d\theta \tag{9}$$

is positive definite. Given such a g, Carathéodory's theorem yields coefficients a_k such that

$$g(\theta)+\mathrm{Re}\sum_s^S a_k e^{ik\theta}>0,$$

which implies (5):

$$\frac{f(\theta)}{(1-\cos\theta)^{s-t}}+\mathrm{Re}\sum_s^S a_k e^{ik\theta}>0 \text{ for } \theta \neq 0 \pmod{2\pi}.$$

Thus the only problem is one of regularization at $\theta = 0$, by constructing g. Consider the truncated function g_n:

$$\begin{cases} g_n(\theta) = 0 \text{ for } |\theta| < 1/n, \\ g_n(\theta) = f(\theta)/(1-\cos\theta)^{s-t} \text{ for } 1/n \leqq |\theta| \leqq \pi. \end{cases}$$

Then we assert that the form (9), with g replaced by g_n, is positive definite for large enough n. Otherwise we should have normalized trigonometric polynomials $P_n(\theta)$ of degree $s-1$ such that

$$\int g_n |P_n|^2 \leqq 0. \tag{10}$$

Some subsequence of the P_n converges to a (normalized) limit P_∞ of degree $s-1$. Since $s>t$, it is easy to see that $P_\infty(0) = 0$; otherwise the left side of (10) would approach $+\infty$, because $f(0)>0$. In fact, the left side will diverge unless $|P_\infty|^2 = (1-\cos\theta)^{s-t}|Q|^2$ for some Q of degree $t-1$. (Thus our assertion is already proved in the case $t = 0$, where degree $(Q) = -1$ implies $Q = 0$, contradicting the normalization of P_∞.)

For arbitrarily large N, we have:

$$\int g_N |P_n|^2 \leqq 0 \text{ for } n \geqq N,$$

by comparison with (10), since $g_N \leqq g_n$. As $n \to \infty$ through the subsequence, we arrive at the following result:

$$\int g_N (1-\cos\theta)^{s-t} |Q|^2 \leqq 0.$$

If now we let $N \to \infty$, we have a contradiction to (8). Therefore (9) is indeed positive definite, if we replace g by g_n with n large enough. Then we may finally choose a trigonometric polynomial g, lying just below g_n, for which (9) remains positive definite. This proves Lemma 2.

3. We can now state, in rather a cumbrous form, the answer to our original question I. Let us suppose that θ^m is the first non-vanishing power in the expansion

$$\operatorname{Re} \sum_0^{r-1} c_j (e^{i\theta}-1)^j = b_m \theta^m + b_{m+1}\theta^{m+1} + \dots . \tag{11}$$

Theorem. *The answer to question* I *is affirmative if and only if the relevant one of the following three conditions is satisfied*:

(1) *If* $m<r$, *then* $m = 2t$ *must be even*, $b_m>0$, *and* $(T_{t-1}(g)u, u)$ *positive definite* (*if* $t>0$), *where* g *is the polynomial*

$$g = \operatorname{Re} \sum_0^{2t-1} c_j (e^{i\theta}-1)^j/(1-\cos\theta)^t;$$

(2) *if* $m \geqq r$ *and* $r = 2s$ *is even, then* $(T_{s-1}(h)u, u)$ *must be positive definite, where*

$$h = \operatorname{Re} \sum_0^{r-1} c_j (e^{i\theta}-1)^j/(1-\cos\theta)^s;$$

(3) *if* $m \geqq r$ *and* $r = 2s-1$, *the form*

$$(T_{s-1}(l)u, u)+\alpha \mid \Sigma u_k \mid^2$$

must be positive definite for large α, *where*

$$l = \mathrm{Re}\, \frac{\sum_0^{r-1} c_j(e^{i\theta}-1)^j + ib_r(-1)^s(e^{i\theta}-1)^r}{(1-\cos\theta)^s}.$$

Proof. (1) $m<r$; Obviously the terms $\sum_r c_j(e^{i\theta}-1)^j$ which we are free to choose in (1) will be $o(\theta^m)$ as $\theta\to 0$. Therefore $\mathrm{Re}\, p(e^{i\theta}) \sim b_m\theta^m$ and we must have $b_m>0$ and $m = 2t$ even, if we are to achieve $\mathrm{Re}\, p(e^{i\theta})>0$ on both sides of $\theta = 0$. Let

$$f(\theta) = \mathrm{Re} \sum_0^{r-1} c_j(e^{i\theta}-1)^j/(1-\cos\theta)^t.$$

Then f is a real trigonometric polynomial with $f(0) = 2^t b_m>0$, so we may apply Lemma 2: (5) can be satisfied if and only if $(T_{t-1}(f)u, u)$ is positive definite. We want to convert this assertion into: condition I can be satisfied if and only if $(T_{t-1}(g)u, u)$ is positive definite.

According to Lemma 1, the real part of $\sum_{2t}^{r-1} c_j(e^{i\theta}-1)^j/(1-\cos\theta)^t$ is the real part of a polynomial of the form $\sum_t^{r-1-t} a_k e^{ik\theta}$. But the first description exactly fits $f-g$. Since a polynomial fitting the second description has no effect on the $(t-1)$-th Toeplitz form,

$$(T_{t-1}(f)u, u) \equiv (T_{t-1}(g)u, u). \tag{12}$$

We pointed out, after the proof of Lemma 1, that satisfying (5) was equivalent to achieving I, when $r = 2s$ is even. Suppose now that $r = 2s-1$; then the answer to I is affirmative if and only if we can prescribe c_{2s-1} in such a way that the resulting problem with $r = 2s$ has an affirmative answer. Since $m<2s-1$, the choice of c_{2s-1} has no effect on the values of m, b_m, or $(T_{t-1}(g)u,u)$. Thus the answer for $r = 2s-1$ is identical with that for $r = 2s$.

(2) $m \geqq r$ and $r = 2s$ even: In this case the reduction from question I, i.e., from (3) to (4), goes through. Furthermore $h(\theta)$, the first term in (7), is a trigonometric polynomial. Therefore we may use Carathéodory's solution directly; the positive definiteness of $(T_{s-1}(h)u, u)$ is the only test.

(3) $m \geqq r$ and $r = 2s-1$: Again the question is whether c_{2s-1} can be prescribed so that the answer with $r = 2s$ becomes affirmative. For the imaginary part of c_{2s-1} we have no option; it must equal the coefficient $(-1)^{s+1}b_r$ which we have put into l, to cancel the coefficient of θ^{2s-1} in $\mathrm{Re} \sum_0^{2s-1} c_j(e^{i\theta}-1)^j$.

Now according to case (2), we have to ask whether the real part A of c_{2s-1} can be chosen to make the form

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\left(l+\frac{A\,\mathrm{Re}\,(e^{i\theta}-1)^{2s-1}}{(1-\cos\theta)^s}\right)\left|\sum_0^{s-1}u_k e^{ik\theta}\right|^2 d\theta \tag{13}$$

positive definite. Given the identity

$$\frac{\mathrm{Re}\,(e^{i\theta}-1)^{2s-1}}{(1-\cos\theta)^s}=\frac{(-2)^s}{2}\sum_{1-s}^{s-1}e^{ij\theta},$$

the second integral in this form is just

$$\frac{\alpha}{2\pi}\int_{-\pi}^{\pi}\sum_{1-s}^{s-1}e^{ij\theta}\left|\sum_0^{s-1}u_k e^{ik\theta}\right|^2 d\theta=\alpha\,|\,\Sigma u_k\,|^2,$$

where $\alpha=(-2)^s A/2$. Thus the answer to I is affirmative if and only if α can be chosen so that the form

$$(T_{s-1}(l)u,\,u)+\alpha\,|\,\Sigma u_k\,|^2 \tag{13'}$$

is positive definite, completing the proof.

All the tests demanded in our Theorem can be carried out on the prescribed coefficients c_j with a fixed number of computations (depending only on r). Question II remains open.

REFERENCES

(1) U. Grenander and G. Szegö, *Toeplitz Forms and their Applications* (University of California Press, Berkeley and Los Angeles, 1958).

(2) G. Strang, Accurate partial difference methods II: Non-linear problems, *Numerische Math.* **6** (1964), 37-46.

(3) G. Strang, Unbalanced polynomials and difference methods for mixed problems, SIAM *J. Numer. Anal.* **2** (1964), 46-51.

(4) G. Strang, Implicit difference methods for initial boundary value problems, *J. Math. Anal. and Applications*, **16** (1966), 188-198.

Massachusetts Institute of Technology.