stichting
mathematisch
centrum

MC

AFDELING MATHEMATISCHE BESLISKUNDE   BW 83/77   NOVEMBER
(DEPARTMENT OF OPERATIONS RESEARCH)

S.H. TIJS

SOME GENERALIZATIONS OF CARATHÉODORY'S THEOREM AND AN APPLICATION IN MATHEMATICAL PROGRAMMING THEORY

Preprint

2e boerhaavestraat 49 amsterdam

Printed at the Mathematical Centre, 49, 2e Boerhaavestraat, Amsterdam.

The Mathematical Centre, founded the 11-th of February 1946, is a non-profit institution aiming at the promotion of pure mathematics and its applications. It is sponsored by the Netherlands Government through the Netherlands Organization for the Advancement of Pure Research (Z.W.O).

AMS(MOS) subject classification scheme (1970): 52A20, 90C05, 60B99

Some generalizations of Carathéodory's theorem and an application in mathematical programming theory*

by

S.H. Tijs

ABSTRACT

In this paper two new generalizations of Carathéodory's theorem are presented. One of these is used in the study of the connection between two related mathematical programming problems. Both theorems extend results of other authors.

KEY WORDS & PHRASES: *Carathéodory's theorem, mathematical programming*

---

* This report will be submitted for publication elsewhere

1. INTRODUCTION

In theorems 1 and 3 of this paper two new generalizations of Carathéodory's theorem are presented. As a simple corollary of theorem 1 we obtain in theorem 2 again a well-known fact. Theorem 3 extends a recent result of COOK [3].

In section 4 of this paper we compare two mathematical programming problems with the aid of theorem 1.

2. NOTATIONS

For a finite set X, $|X|$ is the number of elements of X.

Let V be a subset of $\mathbb{R}^m$. Then the closure of V and the convex hull of V, are denoted by cl(V) and conv(V) respectively. The relative interior of a convex subset G of $\mathbb{R}^m$ is denoted by relint (G).

For each $i \in \{1,2,\dots,m\}$ the map $\pi_i: \mathbb{R}^m \to \mathbb{R}$ is the i-th projection defined by $\pi_i(x_1,x_2,\dots,x_m) := x_i$.

Let μ be a measure on (the Borel subsets of) $\mathbb{R}^m$. Then

$$\text{supp}(\mu) := \{x \in \mathbb{R}^m,\ \mu(U) > 0 \text{ for each open neighbourhood } U \text{ of } x\}.$$

The probability measure with mass 1 in a is denoted by $\varepsilon(a)$.

The set of those infinite sequences $(p_1,p_2,\dots)$ of real numbers, for which $p_i \geq 0$ for each $i \in \mathbb{N}$ and $\sum_{i=1}^{\infty} p_i = 1$, is denoted by S.

$\bar{\mathbb{R}} := \mathbb{R} \cup \{-\infty,\infty\}$.

3. GENERALIZATIONS OF CARATHÉODORY'S THEOREM

First we recall CARATHÉODORY'S THEOREM:
Let V be a subset of $\mathbb{R}^m$ and let $a \in \text{conv}(V)$. Then there exists a finite subset W of V such that $|W| \leq m + 1$ and $a \in \text{conv}(W)$.

For a proof see e.g. [9], p.35.

The following theorem generalises Carathéodory's theorem.

<u>THEOREM 1</u>. *Let* μ *be a probability measure on* $\mathbb{R}^m$ *such that* $\int \pi_i(x)d\mu(x) \in \mathbb{R}$ *for each* $i \in \{1,2,\dots,m\}$. *Let* V *be a subset of* $\mathbb{R}^m$ *with* $\text{supp}(\mu) = \text{cl}(V)$. *Then there exists a finite subset* W *of* V *such that* $|W| \leq m + 1$ *and such*

that the barycenter

$$b(\mu) := \left(\int \pi_1(x)d\mu(x),\dots, \int \pi_m(x)d\mu(x)\right)$$

of μ *is an element of* conv(W).

PROOF. In view of Carathéodory's theorem it is sufficient to show that $b(\mu) \in \text{conv}(V)$. We shall first prove that

(3.1) $\qquad b(\mu) \in \text{relint}(\text{cl}(\text{conv}(V)))$.

Suppose that this is not true. Then we may conclude (cf. theorem (3.3.9) in [9]) that there exists a linear function $f\colon \mathbb{R}^m \to \mathbb{R}$ such that

(3.2) $\qquad f(b(\mu)) \le f(x) \quad$ for each $x \in \text{cl}(\text{conv}(V))$

and such that

(3.3) $\qquad f(b(\mu)) < f(x_0) \quad$ for some $x_0 \in \text{cl}(\text{conv}(V))$.

It follows from 3.3 that there is an $x^* \in V$ such that $f(b(\mu)) < f(x^*)$. Now let $\varepsilon := \frac{1}{2}(f(x^*) - f(b(\mu)))$. Then there exists an open neighbourhood U of x^* such that

(3.4) $\qquad f(x) \ge f(b(\mu)) + \varepsilon \quad$ for each $x \in U$.

Moreover, $\mu(U) > 0$ because $x^* \in \text{supp}(\mu)$. It follows from 3.2 that

(3.5) $\qquad f(x) \ge f(b(\mu)) \quad$ for each $x \in \text{supp}(\mu)$

since $\text{cl}(\text{conv}(V)) \supset \text{cl}(V) = \text{supp}(\mu)$.
But then, in view of 3.4 and 3.5, we have

$$f(b(\mu)) = \int f(x)d\mu(x) \ge f(b(\mu)) + \varepsilon\mu(U) > f(b(\mu))$$

and that is a contradiction. Hence 3.1 holds.
Then, in view of (3.2.13) in [9], we have

$$b(\mu) \in \text{relint}(\text{conv}(V)) \subset \text{conv}(V)$$

and the proof is completed. □

The following theorem is a direct consequence of theorem 1.

THEOREM 2. *Let* $a_0, a_1, a_2, \ldots$ *be an infinite sequence in* $\mathbb{R}^m$ *and* $(q_1, q_2, \ldots) \in S$ *such that* $a_0 = \sum_{j=1}^{\infty} q_j a_j$. *Then there exists an* $r = (r_1, r_2, \ldots) \in S$ *such that at most* $m + 1$ *coordinates of* r *are nonzero and such that* $a_0 = \sum_{j=1}^{\infty} r_j a_j$.

PROOF. Let $P := \{j \in \mathbb{N} ; q_j > 0\}$. Let $V := \{a_j ; j \in P\}$ and $\mu := \sum_{j \in P} q_j \, \varepsilon(a_j)$. Then μ is a probability measure on $\mathbb{R}^m$ such that $b(\mu) = a_0$ and $\text{supp}(\mu) = \text{cl}(V)$. In view of theorem 1 we may conclude that

$$a_0 \in \text{conv}(V) \subset \text{conv}(\{a_1, a_2, \ldots\}),$$

which implies the conclusion of the theorem. □

Theorem 2 has been proved independently by several authors. See e.g. BLACKWELL & GIRSHICK [1], p.48, COOK & WEBSTER [5] and MORSCHE [7].

Applications of theorem 2 in game theory were given by BLACKWELL & GIRSHICK [1], p.50 and TIJS [10], pp.34, 38 and 46.

We now derive an extension of a result of COOK [3] which can also be seen as a generalization of theorem 2.

THEOREM 3. *Let* $D = [d_{ij}]_{i=1,j=1}^{k \quad \infty}$ *be an upper bounded or a lower bounded* $k \times \infty$*-matrix of real numbers and let* $d = (d_1, d_2, \ldots, d_k) \in \mathbb{R}^k$. *Put* $S(D,d) := \{p = (p_1, p_2, \ldots) \in S; Dp \in \bar{\mathbb{R}}^k, Dp \le d\}$. *Let* $x_0, x_1, x_2, \ldots$ *be an infinite sequence in* $\mathbb{R}^m$ *and let* $q = (q_1, q_2, \ldots) \in S(D,d)$ *such that* $x_0 = \sum_{j=1}^{\infty} q_j x_j$. *Then there exists an* $r = (r_1, r_2, \ldots) \in S(D,d)$ *such that at most* $m + k + 1$ *coordinates of* r *are nonzero and such that* $x_0 = \sum_{j=1}^{\infty} r_j x_j$.

PROOF. Note that $Dp \in \mathbb{R}^k$ for each $p \in S(D,d)$ if D is lower bounded, and that $Dp \in (\mathbb{R} \cup \{-\infty\})^k$ if D is upper bounded.

(a) First suppose that $Dq \in \mathbb{R}^k$. Then $(x_0,Dq) = \sum_{j=1}^{\infty} q_j(x_j,D_j)$, where D_j is the j-th column of the matrix D. It follows from theorem 2 (with the (m+k)-dimensional space $\mathbb{R}^m \times \mathbb{R}^k$ in the role of $\mathbb{R}^m$ and (x_0,Dq), (x_1,D_1), $(x_2,D_2),\ldots$ in the roles of $a_0,a_1,a_2,\ldots$) that there is an $r \in S$ with at most $(m+k)+1$ coordinates unequal to zero and such that

$$(x_0,Dq) = \sum_{j=1}^{\infty} r_j(x_j,D_j) = \left(\sum_{j=1}^{\infty} r_jx_j,Dr\right).$$

But then $r \in S(D,d)$ because $Dr = Dq \le d$, and $x_0 = \sum_{j=1}^{\infty} r_jx_j$. Thus we have proved the theorem for the case that $Dq \in \mathbb{R}^k$.

(b) Now suppose that $Dq \notin \mathbb{R}^k$. Then D is upper bounded and so

$$s := \sup\{d_{ij}; i \in \{1,\ldots,k\},\ j \in \mathbb{N}\} \in \mathbb{R}.$$

Further $I := \{i \in \{1,\ldots,k\};\ \sum_{j=1}^{\infty} d_{ij}q_j = -\infty\}$ is a nonempty set. Take a $t \in \mathbb{N}$ such that

$$\sum_{j=1}^{t} d_{ij}q_j \le d_j - \max\{0,s\} \quad \text{for each } i \in I.$$

Let $C = [c_{ij}]_{i=1,j=1}^{k\ \ \infty}$ be the $k \times \infty$-matrix with

$$c_{ij} := \max\{0,d_{ij}\} \quad \text{if } i \in I \text{ and } j > t,$$

and

$$c_{ij} := d_{ij} \quad \text{otherwise.}$$

Put $S(C,d) := \{p \in S;\ Cp \in \bar{\mathbb{R}}^k,\ Cp \le d\}$.
Then it is straightforward to show that $q \in S(C,d)$ and that $Cq \in \mathbb{R}^k$. In view of part (a) of this proof (with C in the role of D) we may conclude that there exists an $r \in S(C,d)$ with at most $m + k + 1$ coordinates unequal to zero such that $x_0 = \sum_{j=1}^{\infty} r_jx_j$. Now $Dr \le Cr \le d$. Hence $r \in S(D,d)$ and we have proved the theorem. □

W.D. Cook proved the above theorem under the two additional assumptions:

(1) The sequence $D_1, D_2, \ldots$ of columns of D is a closed bounded sequence in $\mathbb{R}^k$.

(2) $x_1, x_2, \ldots$ is a closed bounded sequence in $\mathbb{R}^k$.

In his proof he used a duality theorem of semi-infinite programming theory. Our proof is considerably simpler and our result much more general.

Without going into details we note that those results of the paper of COOK, FIELD & KIRBY [4] which were obtained by using Cook's theorem can be strengthened by using theorem 3.

For other generalizations of Carathéodory's theorem we refer to BONNICE & KLEE [2] and REAY [8].

4. AN APPLICATION IN MATHEMATICAL PROGRAMMING THEORY

Let Y be a set and $m \in \mathbb{N}$. Let $f_1, f_2, \ldots, f_m$ be real-valued lower bounded functions on Y, let f_{m+1} be a real-valued bounded function on Y and let $b = (b_1, b_2, \ldots, b_m) \in \mathbb{R}^m$. By $\underline{B}$ we denote the smallest σ-algebra of subsets of Y such that $f_1, f_2, \ldots, f_{m+1}$ are measurable functions. Let R_1 be the family of those finite measures μ on the measurable space $(Y, \underline{B})$ for which

$$(4.0) \qquad \int f_i(y)d\mu(y) \leq b_i \qquad \text{for each } i \in \{1,2,\ldots,m\}.$$

Let C be the convex cone generated by the set of probability measures $\{\varepsilon(y);\ y \in Y\}$, where $\varepsilon(y)$ is the point measure with mass 1 in y. Let R_2 be the subset of those elements μ of C for which (4.0) holds. For $i \in \{1,2\}$ we look at

<u>PROBLEM i</u>. Find the value

$$v_i := \inf_{\mu \in R_i} \int f_{m+1}(y)d\mu(y)$$

and (if possible) an element of the solution set

$$O_i := \{\mu \in R_i;\ \int f_{m+1}(y)d\mu(y) = v_i\}.$$

Note that the problems 1 and 2 coincide if Y is a finite set and that then essentially we have a standard finite linear programming problem.

The following theorem shows that both problems are feasible if one of them is so; that the values of both problems are equal and that both solution sets are nonempty if one of these sets is. Theorem 1 plays a crucial role in the proof of this theorem.

THEOREM 4. *Notations as above. Moreover, let*

$$O_i(\delta) := \{\mu \in R_i;\ \int f_{m+1}(y)d\mu(y) \le v_i + \delta\}$$

for each $i \in \{1,2\}$ *and for each* $\delta \ge 0$. *Then*

(4.1) $\qquad R_1 \neq \emptyset$ *iff* $R_2 \neq \emptyset$

(4.2) $\qquad v_1 = v_2$

(4.3) $\qquad$ *for each* $\delta \ge 0$: $O_1(\delta) \neq \emptyset$ *iff* $O_2(\delta) \neq \emptyset$.

PROOF. Since $R_2 \subset R_1$, we may conclude that

(4.4) $\qquad R_1 \neq \emptyset$ if $R_2 \neq \emptyset$ and $v_1 \le v_2$.

Note that the theorem holds if $R_1 = \emptyset$.

Suppose now that we can show that

(4.5) $\qquad$ for each $\mu \in R_1$, there is a $\tilde{\mu} \in R_2$ such that

$$\int f_i(y)d\mu(y) = \int f_i(y)d\tilde{\mu}(y) \text{ for each } i \in \{1,2,\dots,m+1\}.$$

Then we may conclude that 4.1 holds and that $v_2 \le v_1$, and thus $v_2 = v_1$ in view of 4.4. Furthermore, $\tilde{\mu} \in O_2(\delta)$ if $\mu \in O_1(\delta)$, while it is also obvious that $O_2(\delta) \subset O_1(\delta)$; thus 4.3 holds. Hence, all that remains is the proof of 4.5.

Take $\mu \in R_1$. If $\mu(Y) = 0$, then let $\tilde{\mu} := \mu \in R_2$, and 4.5 holds. Suppose now that $\mu(Y) > 0$. Note that

$$m_i\mu(Y) \leq \int f_i d\mu \leq b_i \quad \text{for each } i \in \{1,\dots,m\}$$

and

$$m_{n+1}\mu(Y) \leq \int f_{m+1}\, d\mu \leq M\mu(Y)$$

where

$$m_i := \inf_{y\in Y} f_i(y) \in \mathbb{R} \text{ for each } i \in \{1,\dots,m+1\} \text{ and } M := \sup_{y\in Y} f_{m+1}(y).$$

Hence $\int f_i\, d\mu \in \mathbb{R}$ for each $i \in \{1,2,\dots,m+1\}$. Let $T\colon Y \to \mathbb{R}^{m+1}$ be the map with $T(y) := (f_1(y),f_2(y),\dots,f_{m+1}(y))$ for each $y \in Y$, and let ν be the probability measure on $\mathbb{R}^{m+1}$ defined by

$$\nu(A) = (\mu(Y))^{-1}\mu(T^{-1}(A)) \quad \text{for each Borel subset } A \text{ of } \mathbb{R}^{m+1}.$$

Then, in view of theorem C in HALMOS [6] p.163, we have

$$\mu(Y) \int \pi_i(x)d\nu(x) = \int \pi_i(T(y))d\mu(y) = \int f_i(y)d\mu(y) \in \mathbb{R}$$

for each $i \in \{1,2,\dots,m+1\}$. Hence

$$b(\nu) := \int x d\nu(x) = (\mu(Y))^{-1} \int T(y)d\mu(y) \in \mathbb{R}^{m+1}.$$

Let $V := T(T^{-1}(\mathrm{supp}(\nu)))$. Then $\mathrm{cl}(V) = \mathrm{supp}(\nu)$. Hence, in view of theorem 1, there exists a finite subset W of V such that $b(\nu) \in \mathrm{conv}(W)$. Let $|W| = k$. Then there are $y_1,y_2,\dots,y_k \in Y$ such that $W =$ $= \{T(y_1),T(y_2),\dots,T(y_k)\}$. Moreover, there is a $(p_1,p_2,\dots,p_k) \in \mathbb{R}^k$, with $p_j \geq 0$ for each $j \in \{1,\dots,k\}$ and $\sum_{j=1}^k p_j = 1$, such that $b(\nu) = \sum_{j=1}^k p_jT(y_j)$. Put $\tilde{\mu} := \mu(Y)\sum_{j=1}^k p_j\, \varepsilon(y_j) \in R_2$. Then

$$\int f_i(y)d\tilde{\mu}(y) = \mu(Y) \sum_{j=1}^{k} p_jf_i(y_j) = \mu(Y)(b(\nu))_i = \int f_i(y)d\mu(y)$$

for each $i \in \{1,2,\dots,m+1\}$ and thus 4.5 holds. □

REFERENCES

[1] BLACKWELL, D. & M.A. GIRSHICK, *Theory of games and statistical decisions*, John Wiley and Sons, Inc., New York (1954).

[2] BONNICE, W. & V.L. KLEE, *The generation of convex hulls*, Math. Ann. 152, 1-29 (1963).

[3] COOK, W.D., *Carathéodory's Theorem with Linear Constraints*, Canadian Math. Bull. 17, 189-191 (1974).

[4] COOK, W.D., C.A. FIELD & M.J.L. KIRBY, *Infinite Linear Programming in Games with Partial Information*, Operations Research 23, 996-1010 (1975).

[5] COOK, W.D. & R.J. WEBSTER, *Carathéodory's Theorem*, Canadian Math. Bull. 15, 293 (1972).

[6] HALMOS, P.R., *Measure Theory*, van Nostrand Comp., Princeton (6th Printing 1959).

[7] MORSCHE, H.G. TER, *Problem* 341, Problem Section, Nieuw Archief voor Wiskunde, 21, 98 (1973).
See also: *Solutions of problem* 341, Nieuw Archief voor Wiskunde 21, 292-293 (1973) and 22, 85 (1974).

[8] REAY, J.R., *Generalizations of a theorem of Carathéodory*, Memoirs Amer. Math. Soc. 54 (1965).

[9] STOER, J. & C. WITZGALL, *Convexity and Optimization in Finite Dimensions* I, Springer-Verlag, Berlin (1970).

[10] TIJS, S.H., *Semi-infinite and infinite matrix games and bimatrix games*, Ph.D. dissertation, Department of Mathematics, Catholic University, Nijmegen, The Netherlands (1975).

Department of Mathematics
Catholic University
Toernooiveld, Nijmegen
The Netherlands
(temporarily at the Mathematisch Centrum, Amsterdam)