ON A CLASS OF PERFECT SETS

P. KESAVA MENON

Let $\{a_n\}$ be a sequence of positive numbers such that

$$\sum_{1}^{\infty} a_n = 1, \tag{1}$$

$$a_n \geqq R_n = \sum_{n+1}^{\infty} a_k \qquad (n = 1, 2, 3, \cdots). \tag{2}$$

Also let S be the set of numbers α defined by

$$\alpha = \sum_{1}^{\infty} a_n' \\ a_n' = 0 \quad \text{or} \quad a_n \qquad (n = 1, 2, 3, \cdots). \tag{3}$$

Then we have the following:

THEOREM 1. *The set S is perfect; two series of the form* (3) *have the same value if and only if they are of the forms*

$$z_1 + z_2 + \cdots + z_{k-1} + 0 + a_{k+1} + a_{k+2} + \cdots, \tag{4}$$

$$z_1 + z_2 + \cdots + z_{k-1} + a_k + 0 + 0 + \cdots \tag{5}$$

where $z_i = 0$ or a_i $(i = 1, 2, \cdots, k-1)$ and $a_k = R_k$; if $a_k > R_k$ then no number of the set S lies between two numbers (4) *and* (5); *every number of the closed interval* (0, 1) *other than those between pairs of numbers of the forms* (4) *and* (5) *belongs to S; if among the relations* (2) *there are an infinity of strict inequalities then the set S is totally disconnected; if all but a finite number of the relations* (2) *are equalities then the set S consists of a finite number of closed intervals; and, finally, the measure of the set S is* $\lim_{n=\infty} 2^n R_n$.

It is obvious that all numbers of the set S lie in the interval (0, 1). It is also obvious that when $a_k = R_k$ the series (4) and (5) have the same value. Let us now suppose that the kth terms are the first which differ in two series of the form (3) having the same value. It is clear that the kth term in one of the series is 0, and in the other a_k. Subtracting the former series from the latter we get

$$a_k - \sum \pm a_{i_n} = 0 \tag{6}$$

Received by the editors October 16, 1947.

where the suffixes i_n run through the whole or a subset of the set of numbers $k+1$, $k+2$, $\cdots$. But from (2) we have

$$a_k \geqq \sum_{k+1}^{\infty} a_n. \tag{7}$$

This can be compatible with (6) only if the suffixes i_n run exactly through the whole set of numbers $k+1$, $k+2$, $\cdots$, all the signs under $\sum$ in (6) are positive, and (7) is an equality. This implies that the two series having the same value are of the forms (4) and (5) with $a_k = R_k$.

If $a_k > R_k$ then the series (5) is obviously greater than (4). Also, of all series of the form (3) whose first k terms are the same as those in (4) the latter is the greatest, and of all series whose first k terms are the same as those in (5) the latter is the least. Hence if the value α of a series lies between (4) and (5) then the first $k-1$ terms of the series cannot all be the same as those of (4) and (5). Let, therefore, the lth term, $l \leqq k-1$, be the first in α to differ from those in (4) and (5). Then α will be greater than (5) if $z_l = 0$ and less than (4) if $z_l = a_l$, which contradicts the assumption that α lies between (4) and (5).

Let us suppose that α is any number in the interval (0, 1) which does not lie between two series of the forms (4) and (5). If α is the sum of a finite number of distinct a_n's then it is obviously a number of the set S. If that is not the case, then let, in the first instance, α lie between 1 and a_1. From (1) it follows that there is a number ν_1 greater than 1 such that

$$a_1 + a_2 + \cdots + a_{\nu_1 - 1} < \alpha < a_1 + a_2 + \cdots + a_{\nu_1}; \tag{8}$$

but since α does not lie between

$$a_1 + a_2 + \cdots + a_{\nu_1} + 0 + 0 + \cdots$$

and

$$a_1 + a_2 + \cdots + a_{\nu_1 - 1} + 0 + a_{\nu_1 + 1} + a_{\nu_1 + 2} + \cdots$$

the right-hand inequality in (8) can be sharpened into

$$\alpha \leqq a_1 + a_2 + \cdots + a_{\nu_1 - 1} + 0 + a_{\nu_1 + 1} + a_{\nu_1 + 2} + \cdots . \tag{9}$$

If (9) is an equality then α is a number of the set S; if it is an inequality then there exists a suffix ν_2 (>2) such that

$$\begin{aligned} &a_1 + \cdots + a_{\nu_1 - 1} + 0 + a_{\nu_1 + 1} + \cdots + a_{\nu_2 - 1} \\ &\qquad < \alpha < a_1 + \cdots + a_{\nu_1 - 1} + 0 + a_{\nu_1 + 1} + \cdots + a_{\nu_2}. \end{aligned} \tag{10}$$

It is, of course, supposed that if $\nu_2=\nu_1+1$, then the first member of the inequalities (10) is just $a_1+\cdots+a_{\nu_1-1}$. As before we may sharpen the right-hand inequality in (10) into

$$\alpha \leqq a_1+\cdots+a_{\nu_1-1}+0+a_{\nu_1+1}+\cdots+a_{\nu_2-1}+0 + a_{\nu_2+1}+a_{\nu_2+2}+\cdots. \tag{11}$$

If (11) is an equality it follows that α is a number of the set S; otherwise, we can find another number ν_3 $(>\nu_2)$ such that

$$\begin{aligned} a_1+\cdots+a_{\nu_1-1}+0+a_{\nu_1+1}+\cdots+a_{\nu_2-1}+0+a_{\nu_2+1}+\cdots+a_{\nu_3-1} \\ <\alpha<a_1+\cdots+a_{\nu_1-1}+0+a_{\nu_1+1}+\cdots+a_{\nu_2-1}+0 \\ +a_{\nu_2+1}+\cdots+a_{\nu_3}. \end{aligned}$$

Proceeding in this manner we either arrive at an expression for α of the form (3) after a finite number of steps, or enclose α between narrower and narrower bounds tending to the common limit α. This common limit is, by the nature of the construction of the bounds, a series of the form (3) so that α is a number of the set S. If, on the other hand, $a_1>\alpha$, then there exists a suffix i (>1) such that

$$a_{i+1}<\alpha<a_i$$

since a_n decreases to zero as n tends to infinity. We proceed by sharpening the right-hand inequality into

$$\alpha \leqq 0+a_{i+1}+a_{i+2}+\cdots$$

and complete the proof that α belongs to S exactly as above.

Since for a given k the number of pairs of numbers of the forms (4) and (5) is finite, the set of all such pairs of numbers is enumerable and therefore the numbers of the interval (0, 1) not contained in S are those of an enumerable set of disjoint open intervals. This completes the proof that S is perfect.

If all but a finite number of the relations (2) are equalities it is obvious that S consists of a finite number of closed intervals. If, on the other hand, an infinity of the relations (2) are inequalities, then the set is totally disconnected. For, let α be any number of the set S, say,

$$\alpha=b_1+b_2+b_3+\cdots, \quad b_n=0 \quad \text{or} \quad a_n \quad (n=1,2,3,\cdots).$$

Given $\epsilon>0$, arbitrarily small, choose n so large that

$$R_n=a_{n+1}+a_{n+2}+\cdots<\epsilon.$$

By hypothesis there exists a number $k>n$ such that $a_k>R_k$. The open interval I whose end points are

$$b_1 + \cdots + b_{k-1} + 0 + a_{k+1} + a_{k+2} + \cdots,$$
$$b_1 + \cdots + b_{k-1} + a_k + 0 + 0 + \cdots$$

is contained in the interval $(\alpha-\epsilon, \alpha)$ or in the interval $(\alpha, \alpha+\epsilon)$ according as $b_k = a_k$ or 0. It follows that there are numbers arbitrarily close to α and not belonging to S which proves that S is totally disconnected.

We may write the relations (2) in the form

$$R_{n-1} \geqq 2R_n \qquad (n = 1, 2, \cdots)$$

from which it follows that

$$1 \geqq 2R_1 \geqq 2^2R_2 \geqq 2^3R_3 \geqq \cdots \tag{12}$$

and hence that $\lim_{n=\infty} 2^nR_n$ exists. We shall now show that this limit is the measure of the set S.

In fact the length of the open interval whose end points are (4) and (5) is obviously

$$a_k - \sum_{k+1}^{\infty} a_n = R_{k-1} - 2R_k;$$

also, for a given k, the number of pairs of numbers of the forms (4) and (5) is 2^{k-1}; therefore the total length of such intervals is

$$2^{k-1}(R_{k-1} - 2R_k);$$

summing over all k from 1 to n we get

$$\sum_{1}^{n} (2^{k-1}R_{k-1} - 2^kR_k) = 1 - 2^nR_n;$$

finally, letting n tend to infinity we get the measure of the complement of the set S from which it follows that the measure of S is $\lim_{n=\infty} 2^nR_n$.

If the relations (2) are all equalities then the relations (12) are also all equalities and we get

$$a_n = R_n = 1/2^n \qquad (n = 1, 2, \cdots).$$

In this case the set S consists of the whole of the interval (0, 1) and the representation of the numbers of S by series of the form (4) is the ordinary binary representation. If the relations (2) are not all equalities then the relations (12) are also not all equalities and so $\lim_{n=\infty} 2^nR_n < 1$.

It is easy to construct sets of the type S having any given measure α, $0\leqq\alpha<1$. In fact, let $\{s_n\}$ be a monotonic decreasing sequence tending to α with $s_0=1$, and let the sequence $\{a_n\}$ be defined by the relations

$$a_n = \frac{s_{n-1}}{2^{n-1}} - \frac{s_n}{2^n} \qquad (n = 1, 2, \cdots).$$

It is clear that the a_n's satisfy conditions (1) and (2) and that the set S constructed with these a_n's has the measure α.

Kakeya has proved[1] that if $\{a_n\}$ is a sequence of positive numbers satisfying the relations

$$\sum_{1}^{\infty} a_n = s, \tag{13}$$

$$a_n \leqq R_n, \qquad a_n \geqq a_{n+1} \qquad (n = 1, 2, 3, \cdots) \tag{14}$$

then the set S of all numbers of the form (3) consists of the whole interval $(0, s)$.

Let us now suppose that $\{a_n\}$ is a sequence of positive numbers satisfying the relations (1) and

$$a_n \geqq R_n \qquad (n = 1, 2, \cdots, k - 1), \tag{15}$$

$$a_n \leqq R_n, \qquad a_n \geqq a_{n+1} \qquad (n = k, k + 1, \cdots), \tag{16}$$

and consider the set S of all numbers of the form (3). If all the relations (15) are equalities then (15) and (16) together reduce to a particular case of Kakeya's relations (14) and the set S will be the whole of the interval $(0, 1)$. Let us therefore suppose that the relations (15) are not all equalities. It can then be shown that S consists of a finite number of disjoint closed intervals.

To prove this let us observe that by Kakeya's Theorem the numbers of the form

$$\sum_{k}^{\infty} a_n', \qquad a_n' = 0 \quad \text{or} \quad a_n,$$

fill the whole interval $I(0, R_k)$. The set S will therefore consist of I and the intervals obtained by moving I towards the right through the distances

$$a_1' + a_2' + \cdots + a_{k-1}', \quad a_i' = 0 \text{ or } a_i \qquad (i = 1, 2, \cdots, k - 1).$$

The interval I and those obtained by the translations have no points

[1] G. Pólya and G. Szegö, *Aufgaben*, Part 1, no. 131.

in common except perhaps end points, and they do not cover the entire interval (0, 1).

In fact, if one of the relations (15) is an inequality, say, $a_l > R_l$, then, as in our previous discussion, it can be seen that there is no number of the set within the interval whose end points are

$$z_1 + \cdots + z_{l-1} + 0 + a_{l+1} + a_{l+2} + \cdots ,$$
$$z_1 + \cdots + z_{l-1} + a_l + 0 + 0 + \cdots .$$

It is also not difficult to see that the only gaps in the set S are those which arise in this manner.

For, if we arrange the numbers

$$a_1' + a_2' + \cdots + a_{k-1}', \qquad a_i' = 0 \text{ or } a_i \, (i = 1, 2, \cdots , k - 1)$$

in increasing order of magnitude, then any consecutive pair will be of the form

$$\alpha = z_1 + z_2 + \cdots + z_{l-1} + 0 + a_{l+1} + a_{l+2} + \cdots + a_{k-1},$$
$$\beta = z_1 + z_2 + \cdots + z_{l-1} + a_l + 0 + 0 + \cdots + 0$$

where $z_i = 0$ or a_i $(i = 1, 2, \cdots , l+1)$, $l \leqq k-1$. Moving the interval I through the distances α and β we get the two intervals $(\alpha, \alpha+R_k)$ and $(\beta, \beta+R_k)$. The right-hand end point of the former is clearly less then or equal to the left-hand end point of the latter according as a_l is greater than or equal to R_l.

As an illustration we may prove the following.

THEOREM 2. *Let $S(t)$ be the set of numbers of the form*

$$t^{i_1} - t^{i_2} + t^{i_3} - \cdots \tag{17}$$

where $\{i_n\}$ is a finite or infinite sequence of increasing positive integers, zero included, and $0 < t < 1$. Then $S(t)$ will consist of the whole of the interval (0, 1) *if $1/2 \leqq t < 1$ and will be a totally disconnected perfect set of measure zero if $0 < t < 1/2$. Also, in the latter case, for a given t, no two series of the form* (17) *represent the same number of the set $S(t)$.*

To prove this we have only to take $a_n = t^{n-1} - t^n$ $(n = 1, 2, \cdots)$. Then for $1/2 \leqq t < 1$ we get relations of the form (14) (Kakeya's type) and, for $0 < t < 1/2$, relations of the form (2). In the latter case it is moreover clear that all the relations are strict inequalities. Hence the first part of the Theorem follows from Kakeya's Theorem and the rest from Theorem 1.

MADRAS CHRISTIAN COLLEGE