PACIFIC JOURNAL OF MATHEMATICS
Vol. 35, No. 2, 1970

PERFECT SUBSETS OF DEFINABLE SETS OF REAL NUMBERS

RICHARD MANSFIELD

There has been some interest in trying to see which sets of real numbers contain perfect subsets. In this paper we prove a theorem which might be able to solve this problem in some generality.

As of the moment however, the only sets to which it applies are the $\sum_2^1$ sets and the $\sum_3^1$ sets. A somewhat weaker version of our specialization to $\sum_2^1$ has been proven by R. M. Solovay [16]. The proof given here (which differs greatly from Solovay's proof) has been known to me for several years but is as yet unpublished. It has the advantage of being a special case of a much more general theorem.

A tree is a set of finite sequences which contains every initial subsequence of each of its members. A sequence t is an extension of the sequence s if both are in the tree and s is an initial subsequence of t; t is an immediate extension if $t \neq s$ and there are no sequences in the tree which are between s and t. A tree has finite branching if every element has only finitely many immediate extensions. For f an arbitrary function with domain the set of integers we let $\bar{f}(n)$ be the finite sequence $\langle f(0), f(1), \cdots, f(n)\rangle$. A path through the tree T is a function f such that for arbitrarily large n $\bar{f}(n)$ is in T. The set of paths through T will be written $[T]$. A closed tree is a tree in which every sequence in the tree has a proper extension. Given an arbitrary set A we can put the discrete topology on A and form the product space A^N where N is the set of nonnegative integers. It is easily seen that any closed subset of A^N can be uniquely written as the set of paths through a closed tree, and conversely that the set of paths through a closed tree is always a closed subset of A^N. Furthermore a closed set will be compact if and only if its tree has finite branching. A perfect set is a closed set with no isolated points. Alternatively, a perfect set is a closed set whose tree has the property that every sequence in the tree has at least two incompatible extensions.

For cartesian products $A^N \times B^N$ we shall use trees of pairs. That is a tree of pairs is a set of pairs of sequences containing the pair $\langle\langle \cdot \rangle, \langle \cdot \rangle\rangle$; the two entries of each pair are required to be of the same length. All the notions and statements of the preceeding paragraph have an obvious analog for trees of pairs. Such a tree is called perfect in the 1^{st} coordinate if every member of the tree has at least two extensions with incompatible 1^{st} coordinates. The following easy proposition contains the combinatorial part of our method.

PROPOSITION 1. *If the tree of pairs T is perfect in the 1^{st} coordinate, then the projection of $[T]$ onto the 1^{st} coordinate contains a perfect subset.*

Proof. This proof is perhaps best seen in the privacy of one's own mind, but at the risk of confusing matters we give it on paper. A function H from the set of finite sequences of 0's and 1's into T is defined by induction. $H(\langle\cdot\rangle) = \langle\langle\cdot\rangle, \langle\cdot\rangle\rangle$ (where $\langle\cdot\rangle$ is the empty sequence). Given $H(s)$, choose $H(s0)$ and $H(s1)$ to be two extensions of $H(s)$ which are incompatible on the 1^{st} coordinate. The set of initial subsequences of the sequences in the range of H is a finite branching tree and so the set of its paths is a compact set. The projection of this compact set onto the 1^{st} coordinate is also a compact set and the incompatibility requirement on $H(s0)$ and $H(s1)$ implies that it has no isolated points.

Now that this rather simple-minded proposition has been proven, we are left with the more difficult task of finding appropriate trees which project into given sets of reals. (The words "real number" are used here in a loose sense to mean a function from N into N. Such functions can be identified with the continued fraction expansions of irrational real numbers.) In order to illustrate the method, let us prove a known theorem.

THEOREM 2. *Every $\sum_1^1$ set of real numbers which contains a nonhyperarithmetic element contains a perfect subset* [2].

LEMMA. *No $\sum_1^1$ set contains exactly one nonhyperarithmetic element.*

Proof. Suppose otherwise that A is a $\sum_1^1$ set and α is its only nonhyperarithmetic element. The set of hyperarithmetic reals is $\prod_1^1$, so $A \sim HYP$ is $\sum_1^1$ and α is its only element. Thus,

$$\begin{aligned} \alpha(n) = m &\equiv \exists\beta[\beta \in A \sim HYP \wedge \beta(n) = m] \\ &\equiv \forall\beta[\beta \in A \sim HYP \rightarrow \beta(n) = m] \end{aligned}$$

so $\alpha \in HYP$.

Proof of Theorem 2. For any $\sum_1^1$ set A there is a recursive tree

T of pairs of sequence of integers such that

$$\begin{aligned}\bar{\alpha} \in A &\equiv \exists\beta\forall n[\langle\alpha(n), \bar{\beta}(n)\rangle \in T] \\ &\equiv \exists\beta[\langle\alpha, \beta\rangle \in [T]] \,.\end{aligned}$$

We define

$$\begin{aligned}T^r = \{\langle s, u\rangle \in T\colon \exists\alpha, \beta[\alpha \notin HYP \wedge \langle s, u\rangle \subset \langle\alpha, \beta\rangle \\ \wedge \langle\alpha, \beta\rangle \in [T]\} \,.\end{aligned}$$

T^r is a subtree of T, so its projection onto the 1^{st} coordinate is a subset of A; the lemma guarantees that it is perfect in the 1^{st} coordinate.

This theorem required us to use only trees of sequences of integers. For what follows we shall use ordinal sequences as the second elements of the sequence pairs. For $\sum_2^1$ sets the ordinals may be all countable, for $\sum_3^1$ it is necessary to use ordinals up to the first cardinal after the first measurable cardinal. If our conjecture turns out to be correct and all definable sets can be analyzed in this way, $\sum_{17}^1$ might require very large ordinals indeed. Or again it might require ordinals no larger than for $\sum_3^1$ but very large cardinal assumptions might be necessary to prove that the process works.

For the next sequence of results we assume that T is a fixed tree of pairs, the first element of each pair consisting of integer sequences and the second of ordinals, and that $\langle s_0, u_0\rangle$ is a fixed element of T.

LEMMA 3. *If $\langle\alpha_n\rangle$ is an infinite sequence of reals and if there is a pair $\langle\alpha, f\rangle$ satisfying the conditions:*

(1) $$\forall n[\alpha_n \neq \alpha]$$

(2) $$\langle s_0, u_0\rangle \subset \langle\alpha, f\rangle$$

(3) $$\langle\alpha, f\rangle \in [T] \,,$$

then there is a pair in $L(\langle\alpha_n\rangle, T)$ satisfying the same conditions.

Proof. Let,

$$\begin{aligned}S = \{\langle s, t\rangle\colon\ & s, t \text{ are sequences of integers} \wedge \text{length } (s) \\ &= \text{length } (t) \wedge \forall n[n < \text{length } (t) \wedge t_n < \text{length } (s) \\ &\rightarrow \alpha_n(t_n) \neq s_{t_n}]\} \,.\end{aligned}$$

Condition (1) is equivalent to

$$\exists\beta[\langle\alpha, \beta\rangle \in [S]] .$$

Let

$$T^1 = \{\langle s, t, u\rangle : \langle s, t\rangle \in S \wedge \langle s, u\rangle \in T \\ \wedge \langle s_0, u_0\rangle \subseteqq \langle s, u\rangle\} .$$

T^1 is in $L(\langle\alpha_n\rangle, T)$ and the existence of an α satisfying conditions (1), (2), and (3) is equivalent to "T^1 is not well founded." But any tree is well-founded if and only if it is well-founded in at least one transitive model for set theory containing it as an element. Thus, in particular, "T^1 is not well-founded" is equivalent to its relativization to $L(\langle\alpha_n\rangle, T)$, completing the proof of the lemma.

Let us now define a condition θ on $L(T)$. We say that $L(T)$ satisfies condition θ if there are two sequences $\langle\alpha_n\rangle$ and $\langle\beta_n\rangle$ each consisting of exactly the reals in $L(T)$ such that all reals that are in both $L(\langle\alpha_n\rangle, T)$ and $L(\langle\beta_n\rangle, T)$ are already in $L(T)$. One instant consequence of θ is that the reals in $L(T)$ are countable; it can be shown that it is in fact equivalent to θ. We are using θ in order to avoid giving that proof.

LEMMA 4. *Condition θ implies that if there is an $\alpha \notin L(T)$ and an f with $\langle s_0, u_0\rangle \subset \langle\alpha, f\rangle$ and $\langle\alpha, f\rangle \in [T]$, then there are at least two such α.*

Proof. Letting $\langle\alpha_n\rangle$ and $\langle\beta_n\rangle$ be as in condition θ, Lemma 3 implies that there is one such α in $L(\langle\alpha_n\rangle, T)$ and another in $L(\langle\beta_n\rangle, T)$.

LEMMA 5. *There is a homogeneous Boolean valued extension of set theory in which condition θ is valid.*

Proof. The following proof relies heavily on standard theorems about Boolean valued set theory. For the basic construction of Boolean valued set theory we refer the reader to Rosser [10]. The actual theorems used can be mainly found in Solovay [18], especially § 2. For the material on Boolean algebras the reader may refer to Halmos [1].

Given any set A there is a canonical Boolean extension of set theory in which A is countable. Namely, we let **B** be Boolean algebra of regular open subsets of the topological space A^N. Then we can introduce a function symbol F and an assignment of truth values to statements of the form $F(n) = a$ in such a way that the Boolean value of the statement "F is a function from N onto A" is 1. The assignment is

$$\| F(n) = a \| = \{f \in A^N : f(n) = a\} .$$

The R.H.S. of this equation is a clopen, hence regular open, subset of A^N and as such is a member of **B**. One easily checks that each clause of the statement " F is a function from N onto A " has value 1.

Now we take $A = L(T) \cap N^N$ and define $\mathbf{C} = \mathbf{B} \oplus \mathbf{B}$ (i.e., **C** is the completion of the direct sum of **B** with itself.). In **C** there are two canonical counting functions for $L(T) \cap N^N$ viz,

$$\| F_1(n) = \alpha \| = \| F(n) = \alpha \| \oplus 1$$

$$\| F_2(n) = \alpha \| = 1 \oplus \| F(n) = \alpha \| .$$

Since **B** is homogeneous one can easily check the **C** is also homo-geneous. So in order to complete the proof of the lemma we must show that in B-valued set theory $L(F_1, T) \cap L(F_2, T) = L(T)$.

$\mathbf{B} \oplus 1$ and $1 \oplus \mathbf{B}$ are both complete subalgebras of **C**. Hence an easy transfinite induction on σ proves that all sets in $L_\sigma(F_1, T)$ take on values only in $\mathbf{B} \oplus 1$ and all sets in $L_\sigma(F_2, T)$ take on values only in $1 \oplus \mathbf{B}$. Hence any set in $L(F_1, T) \cap L(F_2, T)$ takes on values only in $\mathbf{B} \oplus 1 \cap 1 \oplus \mathbf{B} = 2$. Thus any set in both is standard. But now a basic result of Boolean valued set theory says that any standard element of $L(F_i, T)$ is already in $L(T)$ (This is another way of saying that the forcing relation for formulae relativized to $L(F_i, T)$ can be defined in $L(T)$.).

THEOREM 6. *If* $\{\alpha : \exists f[\langle \alpha, f \rangle \in [T]]\}$ *contains an element not in* $L(T)$, *then it contains a perfect subset.*

Proof. Let T^r be the set defined within the Boolean extension of the previous lemma with the formula,

$$\langle s, t \rangle \in T^r \equiv \exists \alpha, f[\langle s, t \rangle \subset \langle \alpha, f \rangle \wedge \langle \alpha, f \rangle \in [T] \wedge \alpha \notin L(T)] .$$

This formula has only standard parameters and hence, since **C** is homogeneous, it takes on only the values 0 and 1. That is to say, even though we used **C**-valued set theory to define T^r, it is a member of the 2-valued universe. In the **C**-valued extension it is a subtree of T and is perfect in the 1^{st} coordinate. But this state-ment can be expressed by a formula of set theory having no un-bounded quantifiers. Being true in **C**-valued set theory it is also true in any transitive model containing all its parameters; hence it is true in the 2-valued universe.

DEFINITION. An S-set is a set of real numbers such that there is a tree of pairs of ordinal sequences T with

$$\alpha \in A \equiv \exists f[\alpha, f\rangle \in [T]] .$$

THEOREM 7. *Any $\sum_2^1$ set with a nonconstructible element contains a perfect subset.*

Proof. J. R. Shoenfield in his proof that any $\sum_2^1$ formula is equivalent to its relativization to L [15], proves that any $\sum_2^1$ set is an S-set with a constructible tree.

In a forthcoming paper entitled "A Souslin operation for $\prod_2^1$," the author proves with the aid of a measurable cardinal that any $\prod_2^1$ set is an S-set with an ordinal definable tree; this construction easily yields that any $\sum_3^1$ set is also an S-set with an ordinal definable tree. Thus,

THEOREM 8. *The existence of a measurable cardinal implies that any $\sum_3^1$ set with a non-ordinal-definable element contains a perfect subset.*

Unfortunately Theorem 8, unlike Theorem 7, is not a best possible result. It is an easy exercise to show that if $V \neq L$ for reals, a $\prod_1^1$ set contains a perfect subset if and only if it not is a subset of L. Moreover if $\aleph_1^{(L)} = \aleph_1$ there is an uncountable $\prod_1^1$ set with no perfect subsets. These results do not seem to be true if we replace $\prod_1^1$ by $\prod_2^1$ and constructible by ordinal definable. However Theorem 8 does allow some flexibility in finding a non-ordinal-definable element of a given $\sum_3^1$ set. For instance it is acceptable to look for this member in any homogeneous mild Cohen extension of set theory.

BIBLIOGRAPHY

1. P. R. Halmos, *Lectures on Boolean Algebras*, Van Nostrand, Princeton, New Jersey, 1963.
2. J. Harrison, doctoral dissertation, Stanford University, Palo Alto, California, 1967.
3. S. C. Kleene, *Arithmetical predicates and function quantifiers*, Trans. Amer. Math. Soc., **79** (1955), 312-340.
4. C. Kuratowski, *Ensembles projectifs et ensembles singuliers*, Fund. Math. **35**(1948), 131-140.
5. ———, *Topologie* v. 1, Panstowe Wydawnictwo Naukowe, Warsaw, 1958.
6. N. Lusin, *Sur l'existence d'un ensemble non-dénombrable qui est de première catégorie sur tout ensemble parfait*, Fund. Math. **2** (1921), 155.
7. R. Mansfield, *A Souslin operation for* Π_2^1, Israel J. Math. (to appear).
8. ———, *The theory of $\sum_2^1$ sets*, doctoral dissertation, Stanford University, 1969.
9. G. Paris, doctoral dissertation, Manchester University, Manchester, England, 1969.
10. J. B. Rosser, *Simplified Independence Proofs*, Academic Press, New York, 1969.

11. R. Salem, *Algebraic Numbers and Fourier Analysis*, Heath, Boston, 1963.
12. D. Scott, *Boolean valued models for set theory*, Proc. Sympos. Pure Math., Vol. **13**, Amer. Math. Soc., Providence, R.I., (to appear).
13. ———, *Measurable cardinals and constructible sets*, Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys. **19** (1961), 521-524.
14. J. R. Shoenfield, *Mathematical Logic*, Addison-Wesley Pub. Co., Menlo Park, California, 1967.
15. ———, *The Problem of Predicativity*, Essays on the Foundations of Mathematics, Magnes Press, Hebrew University, Jerusalem, 1961.
16. R. M. Solovay, *The Cardinality of* $\sum_2^1$ *Sets*, Foundations of Mathematics, Springer-Verlag, New York, 1968.
17. ———, *A non-constructible* Δ_3^1 *set of integers*, Trans. Amer. Math. Soc. **127** (1967), 50-75.
18. ———, *A model for set theory in which every set of reals is Lebesgue measurable* (to appear).

Received October 9, 1969.

UNIVERSITY OF CALIFORNIA LOS ANGELES