Combinatorics, Probability and Computing (2008) 00, 000–000.
DOI: 10.1017/S0000000000000000 Printed in the United Kingdom

Almost odd random sum-free sets

NEIL J. CALKIN[1] and P. J. CAMERON[2]

[1] School Of Mathematics,
Georgia Institute of Technology,
Atlanta,
GA 30332,

[2] School of Mathematical Sciences,
Queen Mary and Westfield College,
Mile End Road,
London, E1 4NS

We show that if S_1 is a strongly complete sum-free set of positive integers, and if S_0 is a finite sum-free set, then with positive probability a random sum-free set U contains S_0 and is contained in $S_0 \cup S_1$. As a corollary we show that with positive probability, 2 is the only even element of a random sum-free set.

1. Introduction

In this paper we shall extend the results of Cameron [5] and Calkin [1] on the structure of a random sum-free set.

A set S of positive integers is *sum-free* if there do not exist $x, y, z \in S$ with $x + y = z$. We shall call a sum-free set ultimately complete if there exists n_0 so that for all $n > n_0$, $n \in S \cup (S + S)$, that is, every sufficiently large integer not in S is a sum of elements in S. We define

$$r_S(n) = |\{x : x \leq n, x, n - x \in S\}|$$

to be the number of distinct representations of n as a sum of elements of S. If

$$\lim_{n \notin S} \frac{r_S(n)}{\log(n)} = \infty$$

then we shall call S *strongly complete*. We note that there are no known examples of sum-free sets for which $r_S(n) \to \infty$ but $r_S(n)/n \to 0$: modular complete sum-free sets give rise to sets for which $r_S(n)$ grows linearly.

Cameron [6] introduced a probability measure μ on the set $\mathcal{S}$ of all sum-free sets as follows: there is a natural bijection from the set $2^{\mathbb{N}}$ to $\mathcal{S}$ which induces a probability measure on $\mathcal{S}$. This measure corresponds to the following construction of a random sum-

free set U:
Set $U = \emptyset$: consider each integer n in order: if $n \in U + U$ then increase n by one: if $n \notin U + U$ then toss a fair coin: if heads, then set $U = U \cup \{n\}$, and increase n by one; otherwise increase n by one.

Observe that if $S \subset \{1, 2, 3, \ldots n\}$ is a finite sum-free set, then

$$\Pr_\mu(U \cap \{1, 2, 3, \ldots n\} = S) = 2^{-n+t}$$

where $t = |(S + S) \cap \{1, 2, 3, \ldots n\}|$, since we have to prescribe the outcome of a cointoss for exactly $n - t$ integers.

2. The Main Result

Cameron [5] showed that if S is the sum-free set corresponding to a complete modular sum-free set (modulo m) then $\Pr(U \subset S) > 0$, and Calkin [1] showed that if S is a strongly complete sum-free set then $\Pr(U \subset S) > 0$. Cameron [6] asked whether the probability that a random sum-free set contains 2 and no other even element is positive. In this paper we prove a much stronger result, replacing 2 by an arbitrary finite sum-free set S_0, and the odd numbers by an arbitrary strongly complete sum-free set S_1.

Theorem 2.1. *Let S_0 be a finite sum-free set, and S_1 be a strongly complete sum-free set: then Pr$(S_0 \subset U \subset S_0 \cup S_1) > 0$.*

In our proof we shall assume that the least element of S_1 is at least twice as large as the largest element of S_0: this is not a severe restriction, since in particular it implies the theorem above.

Proof. Our proof will require a probability measure ν on the set $\mathcal{F}$ of all sum-free sets lying between S_0 and $S_0 \cup S_1$, defined in the following manner: set $U = S_0$, and consider the integers $n \in S_1$ in order: if $n \in U + S_0$, move to the next $n \in S_1$; if $n \notin U + S_0$, toss a coin: if it is heads, then set $U = U \cup \{n\}$ and move to the next $n \in S_1$; otherwise, move to the next $n \in S_1$.

In other words, we randomly construct a sum-free set U constrained to lie between S_0 and $S_0 \cup S_1$: whenever we have a choice of whether to add an element to U we toss a coin to decide. Since the least element of S_1 is greater than twice the largest element of S_0, and since S_1 is sum-free, the only times we have to toss a coin correspond to values in $S_0 + S_1$.

We shall denote by ν_n the measure obtained in this fashion after decisions have been made for all elements less than or equal to n. Then if $\mathcal{F}$ is an event, we define $\mathcal{F}_n = \{F \cap \{1, 2, 3, \ldots, n\} | F \in \mathcal{F}\}$. If $\mathcal{F}$ is the limit of $\mathcal{F}_n$ as $n \to \infty$ (in the sense that $F \in \mathcal{F}$ if and only if $F \cap \{1, 2, 3, \ldots, n\} \in \mathcal{F}_n$ for all n, we have $\nu(\mathcal{F}) = \lim_{n\to\infty} \nu_n(\mathcal{F}_n)$.

In particular, if $\mathcal{F}$ is an event which depends only on elements less than or equal to n, then

$$\nu(\mathcal{F}) = \nu_m(\mathcal{F}_m) \qquad \forall m \geq n$$

since all decisions about elements less than n have been made before the decision about m.

Observe that ν is not just the conditional measure given $S_0 \subset U \subset S_0 \cup S_1$: in the conditional measure, sets for which only a few elements of $I\!N \setminus (S_0 \cup S_1)$ are not sums are weighted more heavily than those having many elements not excluded as sums, since the latter require more coin tosses: with ν this is not the case.

However, the measures μ and ν are related as follows:

Lemma 2.2. *Let $t_n(U) = |\{1, 2, 3, \ldots, n\} \setminus (S_0 \cup S_1 \cup (U + U))|$ be the number of elements of $\{1, 2, 3, \ldots, n\} \setminus (S_0 \cup S_1)$ not represented as a sum in U, that is the number of extra coin-tosses used in the μ model over the ν model. Then*

$$Pr_\mu(S_0 \subset U \subset S_0 \cup S_1) = \lim_{n\to\infty} E_{\nu_n}(2^{-t_n(U)}).$$

Proof. Let $\mathcal{F}$ be the event $S_0 \subset U \subset S_0 \cup S_1$. Then

$$\mathrm{Pr}_\mu(S_0 \subset U \subset S_0 \cup S_1) = \lim_{n\to\infty} \mathrm{Pr}_{\mu_n}(S_0 \subset U \cap \{1, 2, 3, \ldots, n\} \subset S_0 \cup S_1)$$

$$= \lim_{n\to\infty} \sum_{F\in\mathcal{F}_n} \mathrm{Pr}_{\nu_n}(U \cap \{1, 2, 3, \ldots, n\} = F) 2^{-t_n(F)} = \lim_{n\to\infty} \mathrm{E}_{\nu_n}(2^{-t_n(U)}).$$

as claimed. □

Hence, if we wish to show that $\mathrm{Pr}_\mu(S_0 \subset U \subset S_0 \cup S_1) > 0$, it suffices to show that there exists a $c > 0$ so that for all n, $\mathrm{E}_{\nu_n}(2^{-t_n(U)}) > c$ (independently of n).

We shall now show that with positive (ν) probability, $t_n(U)$ is bounded, independently of n; more specifically, we show that if $n \in S_1 + S_1$ then $\mathrm{Pr}_{\nu_n}(n \notin U + U)$ is small; in fact, that

$$\sum_{n\in S_1+S_1} \mathrm{Pr}_{\nu_n}(n \notin U + U) < \infty.$$

Then an effective version of Borel Cantelli will give us our result: indeed, if n_0 is such that

$$\sum_{n\in S_1+S_1, n>n_0} \mathrm{Pr}_{\nu_n}(n \notin U + U) < 1 - \epsilon,$$

then

$$\mathrm{Pr}_{\nu_n}(n \in U + U \quad \forall n \in S_1 + S_1, n > n_0) > \epsilon,$$

and hence

$$\mathrm{E}(2^{-t_n(U)}) > \epsilon 2^{-n_0} > 0,$$

and our proof will be complete.

Let the largest element of S_0 be k, and set $l = \lceil r_S(n)/(2k + 1)\rceil - 1$. Then we have

Lemma 2.3.

$$Pr_{\nu_n}(n \notin U + U) \le \left(1 - 2^{-2(2k+1)}\right)^l$$

Proof. Since we have $r_S(n)$ pairs $x, y \in S$ with $x \le y$, $x + y = n$, we can find $x_1, x_2, x_3, \dots, x_l, y_1, y_2, y_3, \dots y_l$ with $x_i + y_i = n$ and $x_{i+1} - x_i > k$, $y_l - x_l > k$: indeed, just pick every $(k+1)$st pair and discard the pair closest to $n/2$.

The key here is that if we force $x_i - k, x_i - k + 1, \dots, x_i - 2, x_i - 1, x_i + 1, x_i + 2, \dots, x_i + k - 1, x_i + k$ to be omitted from U (requiring at most $2k$ coin tosses to be specified) then the other elements of U have no impact on whether x_i is included in U: moreover, whether or not $x_i \in U$ has no impact on other elements of U.

Now let X_i be 1 if $x_i \in U$ and $x_i - k, x_i - k + 1, \dots, x_i - 2, x_i - 1, x_i + 1, x_i + 2, \dots, x_i + k - 1, x_i + k \notin U$, and 0 otherwise, and define Y_i similarly. Then

$$\Pr_{\nu_n}(X_i = 1 \mid X_1,\ X_2, \dots, X_{i-1},\ X_{i+1}, \dots, X_l,\ Y_1,\ Y_2, \dots Y_l)\ \ge\ 2^{-(2k+1)}$$

and similarly for Y_i. Since $n \in U + U$ can only happen if for each i, at least one of X_i, Y_i is equal to 0, we have

$$\begin{gathered}\Pr_{\nu_n}(X_1Y_1 = 0) \le (1 - 2^{-(2k+1)}) \\ \Pr_{\nu_n}(X_2Y_2 = 0 | X_1Y_1 = 0) \le (1 - 2^{-(2k+1)}) \\ \cdot \\ \cdot \\ \cdot \\ \Pr_{\nu_n}(X_lY_l = 0 | X_1Y_1 = 0, \dots X_{l-1}Y_{l-1} = 0) \le (1 - 2^{-(2k+1)})\end{gathered}$$

and hence

$$\begin{aligned}\Pr_{\nu_n}(n \notin U + U) &\le \Pr_{\nu_n}(X_1Y_1 = 0, X_2Y_2 = 0, \dots, X_lY_l = 0) \\ &\le (1 - 2^{-(2k+1)})^l,\end{aligned}$$

completing the proof of the lemma. □

Since S_1 is strongly complete,

$$\sum_{n \in S_1 + S_1} \left(\left(1 - 2^{-2(2k+1)} \right)^{1/(2k+1)} \right)^{r_S(n)} < \infty$$

and the proof of the theorem is complete. □

We note that everything above is for a fair coin: however, the theorem remains true for a coin with probability p of heads, and $1 - p$ of tails, so long as p is strictly between 0 and 1: we omit the proof, as it is essentially the same as the above.

We also note that the proof of the theorem gives us a way to estimate the probability that $S_0 \subset U \subset S_0 \cup S_1$ rather more effectively than by randomly generating sum-free sets with respect to the measure μ and counting the proportion that have the desired property, namely by generating with respect to the measure ν and estimating the expected value of the random variable $2^{-t_n(U)}$. Computer simulations of this type suggest that the probability that a random sum-free set contains the element 2 and no other even element is about 0.00016.

3. Further Questions

1 It is natural to ask now whether this theorem covers almost all sum-free sets, that is, is it true that with probability 1, a random sum-free set is only finitely far from being contained in a strongly sum-free set?

2 One candidate for showing that the answer to Question 1 is false is the following: for $\alpha \in (0,1) \setminus \mathbb{Q}$, define $S_\alpha = \{n | \{n\alpha\} \in (\frac{1}{3}, \frac{2}{3})\}$ where $\{x\}$ denotes the fractional part of x. Calkin and Erdős [2] have shown that for each irrational α, S_α is incomplete. What is

$$\mathrm{Pr}_\mu(U \subset S_\alpha \text{ for some } \alpha \in (0,1) \setminus \mathbb{Q})?$$

3 An old conjecture of Dickson [7] is equivalent to the following: if S is complete then S is ultimately periodic (i.e. there is a period m and an n_0 so that from n_0, S consists of exactly the same elements modulo m): this would imply that $r_S(n)$ has linear growth or has a bounded subsequence. There is evidence that Dickson's conjecture may be false [3, 4]: if so, do there exist sets with $r_S(n) \to \infty$ but $r_S(n)/n \to 0$?

4 If we construct a random sum-free set using a coin with bias p, we have a new measure $\mathrm{Pr}_{\mu,p}$ on the set of all sum-free sets. Let Odd denote the set of all subsets of the odd numbers: is it true that $\mathrm{Pr}_{\mu,p}(\mathrm{Odd})$ is increasing in p? Given a pair S_0, S_1 of sum-free sets, with S_0 finite and S_1 strongly complete, for which value of p is $\mathrm{Pr}_{\mu,p}(S_0 \subset U \subset S_0 \cup S_1)$ maximized? It is clear that if S_0 is non-empty then the limiting value of this probability as p tends to 0 or 1 is 0 (since if p is small, so is the probability that we include the elements of S_0, and as p tends to 1, the probability that U is contained in the odd numbers tends to 1).

5 It follows from the methods in this paper that, conditioned on the only even element being 2, a random sum-free set almost surely has density 1/6. Moreover, in the case where S_1 comes from a modular complete sum-free set, the limiting density exists and is rational. Is it true that almost surely a random sum-free set (constructed with a fair coin) has a limiting density? If so, must the density be rational?

References

[1] Neil J. Calkin. On the structure of a random sum-free set. Submitted.

[2] Neil J. Calkin and P. Erdős. On a class of aperiodic sum-free sets. To appear, Mathematical Proceedings of the Cambridge Philosophical Society.

[3] Neil J. Calkin and Steven R. Finch. Necessary and sufficient conditions for periodicity of sum-free sets. To appear, Experimental Mathematics.

[4] Neil J. Calkin and Steven R. Finch. Difference densities of sum-free sets. In preparation.

[5] P. J. Cameron. On the structure of a random sum-free set. *Probability Theory and Related Fields*, 76(4):523–531, 1987.

[6] P. J. Cameron. Portrait of a typical sum-free set. In C. Whitehead, editor, *Surveys in Combinatorics 1987*, volume 123 of *London Mathematical Society Lecture Notes*, pages 13–42. Cambridge University Press, 1987.

[7] Richard K. Guy. *Unsolved Problems in Number Theory.* Springer Verlag, 1981. Problem E32.