

On the Structure of a Random Sum-free Set

Peter J. Cameron

School of Mathematical Sciences, Queen Mary College, Mile End Road, London E1 4NS, UK

Summary. There is a natural probability measure on the set Σ of all sum-free sets of natural numbers. If we represent such a set by its characteristic function s , then the zero-one random variables $s(i)$ are far from independent, and we cannot expect a law of large numbers to hold for them. In this paper I conjecture a decomposition of Σ into countably many more tractible pieces (up to a null set). I prove that each piece has positive measure, and show that, within each piece, a random set almost surely has a density which is a fixed rational number depending only on the piece. For example, the first such piece is made up of sets consisting entirely of odd numbers; it has probability 0.218 ..., and its members almost surely have density 1/4.

1. Random Sum-free Sets

A subset S of \mathbb{N} is *sum-free* if, for all $x, y \in S$, $x + y \notin S$. An early occurrence of this concept is in the theorem of Schur [6] asserting that \mathbb{N} cannot be partitioned into finitely many sum-free sets. The analogous result was proved for (arithmetic) progression-free sets by van der Waerden [8]. However, unlike progression-free sets (Szemerédi [7], Furstenberg [3]), sum-free sets can have positive density; the set of all odd numbers is sum-free and has density 1/2. (The *upper* and *lower densities* of a set S are the \limsup and \liminf of $|S \cap \{1, \dots, n\}|/n$ as $n \rightarrow \infty$; if they are equal, the common value is the *density* of S .) The question addressed here is: does a typical sum-free set have a density, and if so, what is it?

There is a natural probability measure on the set Σ of sum-free sets, somewhat analogous to the product measure on the power set of \mathbb{N} . It can be described precisely but informally as follows. Consider the natural numbers in turn, in their usual order. While considering n , if we have already included in S two numbers x and y whose sum is n , then (of course) $n \notin S$; if not, then toss a fair coin to decide whether or not to include n in S .

For subsequent use, I give a more formal definition. First, we represent subsets of \mathbb{N} by their characteristic functions, which are infinite zero-one sequences; this identifies Σ with a subset of $2^{\mathbb{N}}$. Now consider the following

machine M . The input and output of M are elements of $2^{\mathbb{N}}$, called is and os . Among the internal structure of M are variables ip and op , which are regarded as pointers to is and os respectively. The program run by M is as follows:

Initialisation: $ip:=1; op:=1$.
Main loop (repeated infinitely):
 {if (there exist x, y with $x + y = op$ and $os(x) = os(y) = 1$)
 then $os(op):=0$
 else { $os(op):=is(ip); ip:=ip + 1$ };
 $op:=op + 1$ }.

It is readily checked that the function $is \mapsto os$ computed by M is a bijection from $2^{\mathbb{N}}$ to Σ . Now the probability measure on Σ is defined to be the image, under this bijection, of the product measure on $2^{\mathbb{N}}$. Note that the two descriptions agree; for if is records the results of the coin tosses, then os is the characteristic function of the sum-free set produced by the informal description.

We can make some simple statements about upper density:

- Proposition 1.1.** (i) *Any sum-free set has upper density at most 1/2.*
 (ii) *With probability 1, a random sum-free set has upper density at most 1/3.*

Proof. (i) Let S be sum-free and non-empty, and let m be the minimum of S . For any $n > m$, the set

$$(m + S) \cap \{1, \dots, n\}$$

is disjoint from $S \cap \{1, \dots, n\}$, and has cardinality at least $|S \cap \{1, \dots, n\}| - m$. So

$$|S \cap \{1, \dots, n\}| \leq (n + m)/2,$$

from which the result follows.

(ii) Use the notation of M . Suppose that, when $ip = m$, we have $op = n$, and r terms of is up to $is(m)$ are equal to 1, of which the first is in position q . Then there are r terms of os up to $os(n)$ equal to 1, and again the first is in position q . Let $\epsilon > 0$ be given. We may assume that m is sufficiently large that, with probability at least $1 - \epsilon$, q is not too large (say $q < \epsilon m$), and also (by the Law of Large Numbers) $|r/m - 1/2| < \epsilon$. Let S be the sum-free set corresponding to os . As in (i), $(q + S) \cap \{1, \dots, n\}$ is disjoint from $S \cap \{1, \dots, n\}$, and has cardinality at least $r - q$. So $n \geq m + r - q$, and we have

$$\begin{aligned} |S \cap \{1, \dots, n\}|/n &\leq r/(m + r - q) \\ &\leq (1/2 + \epsilon)/(3/2 - 2\epsilon). \end{aligned}$$

The right-hand side tends to 1/3 as $\epsilon \rightarrow 0$.

The conjecture in the next section would have the consequence that 1/3 could be improved to 1/4; the theorem in Sect. 3 shows that 1/4 would be best possible.

Remark 1.2. If it were true that $op \leq c \cdot ip$ for some constant c (for all inputs is), it would follow that a random sum-free set almost surely has positive lower density (at least $1/2c$). It is trivial that $op \leq ip(ip + 1)/2$; but examples derived

from [5] show that we can have $op \geq c \cdot ip^{3/2}$ for some constant c . However, it follows from subsequent results that a linear inequality holds for a subset of positive measure.

We conclude this section with a striking result from [1]. Let Odd be the event $S \subseteq 2\mathbb{N} + 1$, that is, $os(n) = 0$ for all even numbers n .

Proposition 1.3. $0.21759 \dots \leq p(\text{Odd}) \leq 0.21862 \dots$

This result motivates the next section.

2. Modular Sum-free Sets

Let m be a positive integer, and let I_m denote the additive group of integers modulo m . Sum-free subsets of I_m are defined as before; in addition, a subset T of I_m is called *complete* if, for every $z \in I_m \setminus T$, there exist $x, y \in T$ such that $x + y = z$. If T is a sum-free subset of I_m , then $S(m, T)$ denotes the (sum-free) subset of \mathbb{N} consisting of the union of the congruence classes in T , and $E(m, T)$ the event that a random sum-free set is contained in $S(m, T)$. (We call $S(m, T)$ a *modular* sum-free set.) In the notation of the preceding section

$$\text{Odd} = E(2, \{1\}).$$

Note that the representation is not unique; for any $k \geq 1$,

$$S(m, T) = S(km, T + \{0, m, \dots, (k-1)m\}),$$

with a similar equation for events $E(m, T)$. However, each such set or event has a unique “primitive” representation which cannot be written in the form on the right-hand side for any $k > 1$.

Proposition 2.1. (i) *Let T be a sum-free set in I_m which is not complete. Then $p(E(m, T)) = 0$.*

(ii) *Let T_i be complete sum-free sets in $I_{m(i)}$, for $i = 1, 2$, such that $S(m(1), T_1) \neq S(m(2), T_2)$. Then*

$$p(E(m(1), T_1) \wedge E(m(2), T_2)) = 0.$$

Proof. (i) Suppose that $z \in I_m \setminus T$ and z is not expressible as $x + y$ with $x, y \in T$. Then, in the production of any output in $E(m, T)$, all input terms corresponding to output terms $op = km + z$ ($k \in \mathbb{N}$) must be zero; so

$$p(S \cap \{1, \dots, km\} \subseteq S(m, T)) \leq 2^{-k}.$$

(ii) Clearly $E(m(1), T_1) \wedge E(m(2), T_2) = E(m(3), T_3)$ for some sum-free set $T_3 \subseteq I_{m(3)}$ which is not complete.

The main result of this section is a converse to Proposition 2.1(i).

Theorem 2.2. *If T is a complete sum-free set in I_m , then $p(E(m, T)) > 0$; in fact, $p(E(m, T)) \geq (c/2)^{m-|T|}$, where $c = p(E(2, \{1\})) \doteq 0.218$.*

Proof. Our main tool is the following special case of the FKG inequality [2]. Let f and g be monotone increasing real functions on the power set of the finite set S . Then

$$\sum_{X \subseteq S} 1 \sum_{X \subseteq S} f(X) g(X) \geq \sum_{X \subseteq S} f(X) \sum_{X \subseteq S} g(X).$$

It follows immediately that, if f_1, \dots, f_r are monotone increasing functions on k -tuples of subsets of S , where $|S| = n$, then

$$\sum_{X_1, \dots, X_k \subseteq S} \prod_{i=1}^r f_i(X_1, \dots, X_k) \geq 2^{-k(r-1)n} \cdot \prod_{i=1}^r \sum_{X_1, \dots, X_k \subseteq S} f_i(X_1, \dots, X_k).$$

We also require the following result.

Lemma 2.3. *Let $S = \{0, \dots, n-1\}$, and let*

$$F(n) = 2^{-2n} \sum_{X \subseteq S} 2^{|(X+X) \cap S|}$$

and

$$F'(n) = 2^{-3n} \sum_{X, Y \subseteq S} 2^{|(X+Y) \cap S|}.$$

Then both $F(n)$ and $F'(n)$ are decreasing functions, and their limits c and c' are both strictly positive. Moreover, $c < c'$.

Proof. The result for $F(n)$ is proved in [1], where it is established that $c = p(E(2, \{1\}))$ – this will follow from our subsequent discussion. I give the argument for $F'(n)$, which is similar but in some ways easier.

Let $S_\infty(n), S_{ij}(n)$ ($i, j = 0, 1$) be the sets of pairs (X, Y) of subsets of S satisfying the following conditions:

$$\begin{aligned} S_\infty: & n \in X + Y, \\ S_{ij}: & n \notin X + Y, \quad 0 \in X \text{ iff } i = 1, \quad 0 \in Y \text{ iff } j = 1. \end{aligned}$$

Then let

$$f_\sigma = \sum_{(X, Y) \in S_\sigma(n)} 2^{|(X+Y) \cap S|},$$

where σ is one of $\infty, 00, 01, 10, 11$. We have

$$2^{3n} F'(n) = f_\infty + f_{11} + f_{10} + f_{01} + f_{00}.$$

Now each pair (X, Y) has four extensions to a pair of subsets of $\{0, \dots, n\}$ – we can choose whether or not to adjoin n to each set – and for each such pair, the contribution to the sum is either equal to or twice that of (X, Y) . Considering cases, we find that

$$\begin{aligned} 2^{3(n+1)} F'(n+1) &= 8f_\infty + 7f_{11} + 6f_{10} + 6f_{01} + 4f_{00} \\ &= 8 \cdot 2^{3n} F'(n) - (f_{11} + 2f_{10} + 2f_{01} + 4f_{00}), \end{aligned}$$

Thus

$$F'(n+1) = F'(n) - (f_{11} + 2f_{10} + 2f_{01} + 4f_{00})/2^{3(n+1)}.$$

This shows that $F'(n)$ is monotone decreasing. Its limit c' is bounded above by $F'(n)$ for any n . Computation shows that

$$F'(14) = 1296700207278/2^{42} \doteq 0.29484.$$

Now each set $S_{ij}(n)$ has cardinality 3^{n-1} , since for $i=1, \dots, n-1$, one of the four possibilities for membership of i in X and $n-i$ in Y is excluded. Moreover, if $(X, Y) \in S_{10}(n)$ or $(X, Y) \in S_{01}(n)$ then $0 \notin X + Y$, while if $(X, Y) \in S_{00}(n)$ then $0, 1 \notin X + Y$. So for $n \geq 2$,

$$f_{11} + 2f_{10} + 2f_{01} + 4f_{00} \leq 4 \cdot 3^{n-1} \cdot 2^n,$$

and so

$$F'(n+1) \geq F'(n) - 3^{n-1}/2^{2n+1}.$$

Thus

$$\begin{aligned} c' &\geq F'(n) - \sum_{i=n}^{\infty} 3^{i-1}/2^{2i+1} \\ &= F'(n) - 3^{n-1}/2^{2n-1} \end{aligned}$$

for $n \geq 2$.

One sees easily that $F'(2) = 15/32 > 3/8$, establishing the positivity of c . From the computed value of $F'(14)$, we have

$$0.28295 \dots \leq c' \leq 0.29484 \dots$$

Proof of Theorem 2.2. Let $|T|=k$ and $T = \{t_1, \dots, t_k\}$ with $0 < t_i < m$ for all i . We have

$$p(E(m, T)) = \lim_{n \rightarrow \infty} P_n,$$

where

$$\begin{aligned} P_n &= p(S \cap \{1, \dots, mn\} \pmod m \subseteq T) \\ &= 2^{-nm} \sum_{\substack{S \subseteq \{1, \dots, nm\} \\ S \pmod m \subseteq T}} 2^{|(S+S) \cap \{1, \dots, nm\}|}. \end{aligned} \tag{*}$$

Let $X_i = \{j | 0 \leq j \leq n-1, mj + t_i \in S\}$ for $i=1, \dots, k$. Then there is a bijection between the sets S in (*) and the k -tuples (X_1, \dots, X_k) of subsets of $\{0, \dots, n-1\}$.

For each $h \in \{1, \dots, m\} \setminus T$, choose $i(h), j(h)$ such that $t_{i(h)} + t_{j(h)} = h$ or $h+m$. Now the number of members of $(S+S) \cap \{1, \dots, nm\}$ which are congruent to $h \pmod m$ is at least $|X_{i(h)} + X_{j(h)} \cap \{0, \dots, n-1\}|$ if $t_{i(h)} + t_{j(h)} = m$, or at least this number minus one if $t_{i(h)} + t_{j(h)} = h+m$. So

$$P_n \geq 2^{-nm} \cdot 2^{-(m-k)} \sum_{X_1, \dots, X_k \subseteq \{0, \dots, n-1\}} \prod_{h \notin T} 2^{|(X_{i(h)} + X_{j(h)}) \cap \{0, \dots, n-1\}|}.$$

The summands are obviously monotone increasing; so the FKG inequality gives

$$P_n \geq 2^{-nm} \cdot 2^{-(m-k)} \cdot 2^{-k(m-k-1)n} \prod_{h \notin T} \sum_{X_1, \dots, X_k \subseteq \{0, \dots, n-1\}} 2^{|(X_{i(h)} + X_{j(h)}) \cap \{0, \dots, n-1\}|}.$$

Now, by the lemma, the value of the sum is at least $2^{(k+1)n} c$ if $i(h)=j(h)$, and at least $2^{(k+1)n} c'$ otherwise. Since $c < c'$ we have

$$P_n \geq 2^{-mn} \cdot 2^{-(m-k)} \cdot 2^{-k(m-k-1)n} \cdot 2^{(k+1)(m-k)n} \cdot c^{m-k} \\ = (c/2)^{m-k},$$

since $-m-k(m-k-1)+(k+1)(m-k)=0$.

Conjecture 2.4. With probability 1, a random sum-free set is contained in some complete modular sum-free set. Equivalently (by Proposition 2.1), if $S(m(1), T_1), S(m(2), T_2), \dots$ is a list of all complete modular sum-free sets, then $\sum p(E(m(i), T_i))=1$.

Remark 2.5. It is unlikely that the above methods will prove the conjecture. For example, $p(E(5, \{2, 3\})) \geq (c/2)^3$. Examining the proof more closely, we see that in fact $p(E(5, \{2, 3\})) \geq c^2 c'/2 \doteq 0.007$. However, computation suggests that the true value is about 0.022.

Remark 2.6. It is easily shown that primitive complete sum-free sets mod m exist for all m except 1, 3, 4, 6, 7, 9, 10 and 15. All such sets have been determined for $m \leq 36$ by N. Calkin (personal communication), and many large examples have been constructed, using a randomised algorithm, by D.J.A. Welsh (personal communication).

3. A Law of Large Numbers

The main result of this section is the following.

Theorem 3.1. *Let T be a complete sum-free set in I_m . Then, conditioned on $E(m, T)$, a random sum-free set almost surely has density $|T|/2m$.*

In other words, the subset of $E(m, T)$ consisting of sets for which the density either fails to exist or takes a value different from $|T|/2m$ is null. In particular, a positive proportion of sum-free sets have density equal to $1/4$.

Proof. We require some terminology. Let k be the function enumerating $S(m, T)$ in order; let X_n be the random variable $os(k(n))$ on the space $E(m, T)$ with the conditional probability induced from Σ (that is, for any event E , the probability of E is

$$p'(E) = p(E \wedge E(m, T))/p(E(m, T)).$$

Let $S_n = \sum_{i=1}^n X_i$. Clearly it suffices to show that

$$S_n/n \rightarrow 1/2$$

almost surely. Of course, this would be true if the X_n were independent with probability $1/2$; this is not the case, but our proof is based on the fact that it is “almost true”.

Before beginning, however, we give one example to show that the statement is not absolutely true. Consider the case when $m=2, T=\{1\}, E(m, T)=\text{Odd}$.

Of course, $p(os(1)=1)=1/2$; but it is an exercise to show that $p(os(1)=1|Odd)=1/8$. (Informally, a head on the first toss increases the likelihood of Odd by a factor of 7/4, while a tail decreases it by a factor of 1/4.) Furthermore,

$$p(os(1)=os(3)=\dots=os(2n-1)=0|Odd)=(1/8)^n;$$

but $p(os(3)=0|Odd)\neq 1/8$, so X_1 and X_2 are not independent in Odd.

The facts we require are:

Lemma 3.2. (i) $p(S \cap \{1, \dots, n\} \subseteq S(m, T)) \leq p(E(m, T)) + c_1^n$ for some $c_1 < 1$.

(ii) $|p'(X_n=1) - 1/2| = O(c_2^n)$ for some $c_2 < 1$.

(iii) For any $(\epsilon_1, \dots, \epsilon_n) \in \{0, 1\}^n$, $p'(X_1=\epsilon_1, \dots, X_n=\epsilon_n) \leq c_3 \cdot 2^{-n}$ for some c_3 .

Proof. (i) Set $E = E(m, T)$, and $E(n)$ denote the event $S \cap \{1, \dots, n\} \subseteq E(m, T)$. Then $\bigwedge E(n) = E$ and $E(n) \supseteq E(n+1)$; so we have

$$\begin{aligned} p(E(n)) - p(E) &= \sum_{i=n}^{\infty} p(E(i) \setminus E(i+1)) \\ &= \sum_{i=n}^{\infty} p(E(i) \wedge (i+1 \notin S(m, T)) \wedge os(i+1)=1). \end{aligned}$$

Thus we must estimate the i th summand on the right. Clearly it is 0 if $i+1 \in S(m, T)$, so suppose not. Then, by the completeness of T (Proposition 2.1), there exist $x, y \in T$ such that $x+y \equiv i+1 \pmod{m}$. (We assume that x and y are integers between 0 and $m-1$.) If $i+1 = x+y+km$, and $os(i+1)=1$, then at most one of each pair $os(x+jm)$, $os(y+(k-j)m)$ is equal to 1 for $j=0, \dots, k$. Let $|S(m, T) \cap \{1, \dots, i\}| = r$. Then each initial sequence of length n in $E(m, T)$ has probability at most 2^{-r} (since all terms with indices in $S(m, T)$, and perhaps others, are determined by coin tosses); and, of the 2^r such sequences, at most $(3/4)^{\frac{1}{2}k} \cdot 2^r$ satisfy the above requirement. So the i th term on the right of the summation has probability at most $(\sqrt{3}/2)^k$, and $k \geq [(i+1)/m] - 1$. Thus the series is dominated by a geometric progression with common ratio $(\sqrt{3}/2)^{1/m} < 1$, and the conclusion follows.

(ii) Clearly, if $E(k(n)-1)$ holds, then $k(n)$ is not the sum of two smaller numbers x and y with $os(x)=os(y)=1$; so

$$p(E(k(n)-1) \wedge (os(k(n))=1)) = \frac{1}{2} p(E(k(n)-1)).$$

Also, $p(E) \leq p(E(k(n)-1)) \leq p(E) + c_1^{k(n)-1}$, so

$$p(E \wedge (os(k(n))=1)) \geq p(E) - \frac{1}{2} p(E(k(n)-1)),$$

and so

$$|p(E \wedge (os(k(n))=1)) - \frac{1}{2} p(E)| \leq \frac{1}{2} c_1^{k(n)-1}$$

Thus $p'(X_n=1) = p(E \wedge (os(k(n))=1))/p(E)$ satisfies

$$|p'(X_n=1) - \frac{1}{2}| = O(c_1^{k(n)}).$$

Since $k(n)$ is bounded below by a linear function of n , we have the result.

(iii) The probability we must estimate is

$$p(E \wedge (os(k(1)) = \varepsilon_1) \wedge \dots \wedge (os(k(n)) = \varepsilon_n) / p(E).$$

The numerator is no greater than

$$p((os k(1)) = \varepsilon_1) \wedge \dots \wedge (os(k(n)) = \varepsilon_n) \wedge (os(r) = 0 \text{ for all other } r \leq k(n)),$$

which does not exceed 2^{-n} , since this event requires favourable results of coin tosses at least for positions $k(1), \dots, k(n)$ in the output sequence. The result follows, with $c_3 = 1/p(E)$.

Remark 3.3. The proof of Proposition 1.3 is obtained by combining the bound for $p(E(n)) - p(E)$ in Lemma 3.2(ii) with a computed value of $p(E(n))$, in the case where $E = \text{Odd}$. The bounds quoted involve computation for $n = 66$ together with some refinements to the estimates.

The completion of the proof of Theorem 3.1 is more or less standard, following the proof of the usual Strong Law of Large Numbers (see [4] Theorem 7.4.3). Let $U_n = |S_n - \frac{1}{2}n|$. We see that $U_{n+1} = U_n + \frac{1}{2}$ if either $U_n = 0$ or $X_{n+1} - \frac{1}{2}$ has the same sign as U_n , while $U_{n+1} = U_n - \frac{1}{2}$ otherwise. Now

$$\begin{aligned} p(U_n = 0) &\leq \binom{n}{\frac{1}{2}n} c_2 \cdot 2^{-n} \quad (\text{by 3.2(iii)}) \\ &= O(n^{-\frac{3}{2}}); \end{aligned}$$

so

$$E(U_{n+1}) = E(U_n) + O(c_1^n) + O(n^{-\frac{3}{2}})$$

by 3.2(ii), and the term $O(c_1^n)$ can be neglected. Summing, we obtain

$$E(U_n) = O(n^{\frac{1}{2}}).$$

Thus

$$\begin{aligned} p(|S_n/n - \frac{1}{2}| > \varepsilon) &= p(U_n > \varepsilon n) \\ &= O(n^{-\frac{1}{2}}/\varepsilon). \end{aligned}$$

Since $\sum n^{-3/2}$ converges, the subsequence (S_{n^3}/n^3) converges almost surely to $\frac{1}{2}$. (See [4], Theorem 7.2.4(c).) Now, if $p^3 \leq n \leq (p+1)^3$, we have $S_{p^3} \leq S_n \leq S_{(p+1)^3}$, so

$$S_{p^3}/(p+1)^3 \leq S_n/n \leq S_{(p+1)^3}/p^3;$$

and since $(p+1)^3/p^3$ and $p^3/(p+1)^3$ tend to 1 as $p \rightarrow \infty$, $S_n/n \rightarrow \frac{1}{2}$ almost surely.

Remark 3.4. It is possible to extract from this proof a sufficient condition for a sequence of zero-one variables to satisfy the Strong Law of Large Numbers: we require that $p(X_n = 1)$ tends, not too slowly, to $\frac{1}{2}$, and that for any $(\varepsilon_1, \dots, \varepsilon_n) \in \{0, 1\}^n$, the probability that $X_i = \varepsilon_i$ for $i = 1, \dots, n$ is not too much greater than 2^{-n} . Further generalisations are also possible.

4. Conclusion

Theorems 2.2 and 3.1 show that there is a countable “spectrum” of densities of sum-free sets which occur with positive probabilities; moreover, if Conjecture 2.4 is true, then almost all sum-free sets have a density in this spectrum. The spectrum consists of all numbers $|T|/2m$, where T is a complete sum-free set mod m . Let \mathcal{S} be the set of such numbers.

Calkin’s (unpublished) catalogue of modular complete sum-free sets suggests a plausible conjecture about \mathcal{S} . Note that, for every k , the set $k+1, k+2, \dots, 2k+1$ is complete sum-free in I_{3k+2} , contributing the value $(k+1)/2(3k+2)$ to \mathcal{S} . Note, also, that if T is complete sum-free in I_m , then so is dT , where $(d, m) = 1$; call such sets *equivalent*.

Conjecture 4.1. A primitive complete sum-free set T mod m with $|T|/m > 1/3$ is equivalent to $\{k+1, k+2, \dots, 2k+1\}$, with $m = 3k+2$, for some k .

This would imply that the largest limit point of \mathcal{S} is $1/6$. Various other limit points are known, including 0 (Hanson and Seyffarth [5]).

References

1. Cameron, P.J.: Cyclic automorphisms of a countable graph and random sum-free sets. *Graphs Comb.* **1**, 129–135 (1985)
2. Fortuin, C.M., Kasteleyn, P.W., Ginibre, J.: Correlation inequalities on some partially ordered sets. *Commun. Math. Phys.* **22**, 89–103 (1971)
3. Furstenberg, H.: Ergodic behaviour of diagonal measures and a theorem of Szemerédi on arithmetic progressions. *J. Anal. Math.* **31**, 204–256 (1977)
4. Grimmett, G., Stirzaker, D.: *Probability and Random Processes*. Oxford: Clarendon Press 1982
5. Hanson, D., Seyffarth, K.: k -saturated graphs of prescribed minimal degree. *Congr. Numerantium* **42**, 169–182 (1984)
6. Schur, I.: Über die Kongruenz $x^m + y^m \equiv z^m \pmod{p}$. *Jber. Deutsch. Math.-Verein.* **25**, 114–116 (1916)
7. Szemerédi, E.: On sets of integers containing no k elements in arithmetic progression. *Acta Arith.* **27**, 199–245 (1975)
8. van der Waerden, B.L.: Beweis einer Baudetschen Vermutung. *Nieuw Arch. Wiskd., III. Ser.* **15**, 212–216 (1927)

Received September 29, 1986