THE DENSITY OF THE PRODUCT OF ARITHMETIC PROGRESSION

S. K. STEIN
University of California, Davis, California

This paper is devoted to the proof of the following theorem.

Theorem. Let a, b, c, and d be positive integers such that $(a,b) = 1 = (c,d)$, Then the density of numbers of the form

$$(ax + b)(cy + d) ,$$

where x and y range over the positive integers, is

$$\frac{1}{(a,c)} .$$

The question arose in the study of the density of products of sets of integers.

The proof is elementary except for the use of Dirichlet's theorem on primes in an arithmetical progression.

1. INTRODUCTION

Let A be a set of positive integers. For a positive integer n let $A(n)$ denote the number of elements in A that lie in the interval $[1,n]$. The upper density of A, $\bar{\delta}(A)$, is defined as

$$\lim_{n\to\infty} \sup A(n)/n .$$

Similarly, the lower density of A, $\underline{\delta}(A)$, is defined as

$$\lim_{n\to\infty} \inf A(n)/n .$$

If $\underline{\delta}(A) = \bar{\delta}(A)$, A is said to have a density, $\delta(A)$, which is the common value of its lower and upper densities. For instance, the arithmetic progression $ax + b$ has density $1/a$.

Consider, as another example, the set

$$S = \{(2x + 1)(2y + 1) \mid x, y \geq 1\} ,$$

which we will abbreviate to

$$\{(2x + 1)(2y + 1)\} .$$

S is clearly contained in the set of odd numbers, $\{2t + 1\}$, whose density is $1/2$. However, S does not exhaust this progression, since no prime is in S. But since the set of primes has density 0, S has density $1/2$, the same density as the set $\{2t + 1\}$. This sets the stage for the following concept, which will be used often in the proof of the main theorem.

Definition. Let a, b, c, and d be positive integers. The set

$$S = \{(ax + b)(cy + d)\}$$

is full if it has the same density as the arithmetic progression

$$\{(a,c)t + bd\} ,$$

namely $1/(a,c)$.

Since S lies in $\{(a,c)t + bd\}$, the definition simply asserts that S fills the progression except for a set of density 0.

The proof of the main theorem depends primarily on the following lemma.

Lemma 2.3. Let a, b, and d be positive integers such that $(a,b) = 1$. Then the set $\{(ax + b)(ay + d)\}$ is full, that is, has density $1/a$.

The general outline of the proof of the theorem is illustrated by the following example, which will be helpful as a reference point when following the proof.

Say that we wish to prove that $\{(2x + 1)(9y + 1)\}$ is full, that is, has density $1/(2,9) = 1$. We begin as follows. The progression $\{2x + 1\}$ is the disjoint union of the nine progressions

$$\{18x + 1\},\ \{18x + 3\},\ \cdots,\ \{18x + 1 + 2i\},\ \cdots,\ \{18x + 17\} .$$

The progression $\{9y + 1\}$ is the disjoint union of the two progressions

$$\{18y + 1\} \qquad \text{and} \qquad \{18y + 10\} .$$

Consequently, $\{(2x + 1)(9y + 1)\}$ is the union of the eighteen sets

$$\left\{\{(18x + 1 + 2i)(18y + 1 + 9j)\} \,\middle|\, 0 \le i \le 8,\quad 0 \le j \le 1\right\} .$$

It is not hard to show that these eighteen sets are disjoint. If we showed that each is full, that is, has density $1/18$, we would be done, for then $\{(2x + 1)(9y + 1)\}$ would have density $18/18 = 1$. Lemma 2.3 shows that fifteen of the eighteen sets are full. Three cases remain, namely

$$\text{(i)} \qquad \{(18x + 3)(18y + 10)\}$$

$$\text{(ii)} \qquad \{(18x + 9)(18y + 10)\}$$

$$\text{(iii)} \qquad \{(18x + 15)(18y + 10)\} .$$

Let $\underline{\delta}$ and $\overline{\delta}$ denote the lower and upper densities of $\{(2x+1)(9y+1)\}$. Clearly $\underline{\delta} \leq \overline{\delta} \leq 1$. We know at this point that

$$1 - 3/18 \leq \underline{\delta} .$$

This completes the first stage.

The second stage consists of repeating the argument of the first stage on each of the unsettled cases (i), (ii), and (iii).

Analysis for case (i): Write $(18x+3)(18y+10)$ as

$$6(6x+1)(9y+5) ,$$

and treat $\{(6x+1)(9y+5)\}$ by the method already illustrated. It turns out that $\{(6x+1)\times(9y+5)\}$ is the disjoint union of the six sets

$$\{\{(18x+1+6i)(18y+5+9j)\} \mid 0 \leq i \leq 2, \ 0 \leq j \leq 1\} .$$

Each is covered by Lemma 2.3, hence has density $1/18$. Thus the density of $\{(18x+3)\times(18y+10)\}$ is

$$\frac{1}{6} \cdot \frac{6}{18} = \frac{1}{18} .$$

(Note that $\{(18x+3)(18y+10)\}$ is full, even though not covered by Lemma 2.3.)

Analysis for case (ii): Write $(18x+9)(18y+10)$ as

$$18(2x+1)(9y+5) ,$$

and apply the argument of the first stage to

$$\{(2x+1)(9y+5)\} . \tag{1.1}$$

It turns out that (1.1) is the disjoint union of the eighteen sets

$$\{(18x+1+2i)(18y+5+9j) \mid 0 \leq i \leq 8, \ 0 \leq j \leq 1\} .$$

Fifteen are covered by Lemma 2.3 and have a total density of $15/18$. Three cases remain,

$$\{(18x+3)(18y+14)\}, \quad \{(18x+9)(18y+14)\}, \quad \{(18x+15)(18y+14)\} .$$

Analysis for case (iii): Write $(18x+15)(18y+10)$ as

$$6(6x+5)(9y+5) .$$

It turns out that $\{(6x + 5)(9y + 5)\}$ splits into six cases, each covered by Lemma 2.3, and case (iii) has density $(1/6)(6/18) = 1/18$.

Combining these three cases, we find at the end of the second stage that

$$1 - 3/18^2 \le \underline{\delta} \le \overline{\delta} \le 1 .$$

Note that the sets removed at each stage by Lemma 2.3 are full.

It turns out that as the process is continued through the third and higher stages, Lemma 2.3 disposes of some sets, and they are full, while the amount left unsettled at the n^{th} stage diminishes toward 0 as $n \to \infty$.

The proof of the theorem shows that the phenomena exhibited by this example occur in general.

2. LEMMAS

We shall make use of the following two basic number-theoretic lemmas. The notation is in the form in which it is used in the proof of the theorem.

Lemma 2.1. Let a, B, c, and D be positive integers such that $(a, B) = 1 = (c, D)$. Then

$$\left[\frac{[a,c]}{(c,B)}, \frac{[a,c]}{(a,D)}\right] = [a,c] .$$

To prove this, consider first the case where a and c are powers of the same prime p, $a = p^{a'}$ and $c = p^{c'}$. The general case follows immediately by using the prime factorization of a and c.

Lemma 2.2. Let a, b, c, and d be positive integers such that $(a,b) = 1 = (c,d)$. Then the sets

$$\{([a,c]\,x + b + ia)([a,c]\,y + d + jc)\} ,$$

$$0 \le i \le ([a,c]/a) - 1, \qquad 0 \le j \le ([a,c]/c) - 1 ,$$

are disjoint.

Proof. Assume that

$$(b + ia)(d + jc) \equiv (b + i'a)(d + j'c) \pmod{[a,c]} ,$$

or equivalently,

$$(i - i')ad \equiv (j' - j)bc \pmod{[a,c]} .$$

Then

$$(i - i')ad \equiv (j' - j)bc \pmod{c/(a,c)} ,$$

hence

$$(i - i')ad \equiv 0 \pmod{c/(a,c)} .$$

Since ad and $c/(a,c)$ are relatively prime, it follows that

$$i - i' \equiv 0 \pmod{c/(a,c)} .$$

But

$$0 \le |i - i'| \le ([a,c]/a) - 1.$$

Since

$$[a,c]/a = c/(a,c) ,$$

it follows that $i = i'$. Similarly, $j = j'$, and the lemma is proved.

The next lemma depends on certain properties of primes that we now review. If $p_1, p_2, \cdots, p_k$ are distinct primes, then the set of positive integers divisible by none of them has density

$$\prod_{i=1}^{k} (1 - p_i^{-1}) \quad .$$

Consequently, if $p_1, p_2, \cdots, p_i, \cdots$ is an infinite sequence of primes such that

$$\sum p_i^{-1} = \infty ,$$

then the set of positive integers divisible by at least one of the p_i's has density 1.

Dirichlet's theorem on primes in arithmetic progressions implies that if a and b are relatively prime positive integers, then the arithmetic progression $\{ax + b\}$ contains an infinite set of primes, $p_1, p_2, \cdots, p_i, \cdots$, and $\sum p_i^{-1} = \infty$.

With this background we are ready for the main lemma.

Lemma 2.3. Let a, b, and d be positive integers such that $(a,b) = 1$. Then the set

$$\{(ax + b)(ax + d)\}$$

is full, that is, has density $1/a$.

Proof. Consider the set $\{az + bd\}$. By Dirichlet's theorem, the density of the subset of $\{(ax + b)(ay + b)\}$ that is divisible by at least one prime p of the form $ax + b$ is $1/a$. If $n = az + bd$ is divisible by p, there is an integer q such that

$$n = az + bd = pq .$$

Taking congruences modulo a we have

$$bd \equiv pq \pmod{a} .$$

Hence

$$bd \equiv bq \pmod{a} ,$$

and since $(a,b) = 1$,

$$d \equiv q \pmod{a} .$$

Thus q has the form $ay + d$, and we conclude that $n = az + bd$ is an element of $\{(ax + b)\times(ay + d)\}$. This proves the lemma.

3. PROOF OF THE THEOREM

Let S be the set

$$\{(ax + b)(cy + d)\} \quad , \tag{3.1}$$

where $(a,b) = 1 = (c,d)$. We wish to prove that S is full. To begin, let $\underline{\delta}$ and $\overline{\delta}$ be the lower and upper densities of (3.1). Clearly $\overline{\delta} \le 1/(a,c)$. We will now show that $\underline{\delta} = 1/(a,c)$.

The progression $\{ax + b\}$ is the disjoint union of the $[a,c]/a$ progressions

$$\{[a,c]\, x + b + ia\}, \quad 0 \le i \le ([a,c]/a) - 1.$$

Similarly, $\{ay + d\}$ is the disjoint union of the $[a,c]/c$ progressions

$$\{[a,c]\, y + d + jc\}, \quad 0 \le j \le ([a,c]/c) - 1 .$$

Thus, by Lemma 2.2, S is the disjoint union of the

$$\frac{[a,c]}{c} \cdot \frac{[a,c]}{a}$$

sets

$$\{([a,c]\, x + b + ia)([a,c]\, y + d + jc)\} , \tag{3.2}$$

$$0 \le i \le ([a,c]/a) - 1, \quad 0 \le j \le ([a,c]/c) - 1 .$$

For convenience, let $B = b + ia$ and $D = d + jc$ in (3.2). Note that $(a,B) = 1 = (c,D)$. Some of the sets (3.2) are covered by Lemma 2.3. Let us call them "good." Some may not be and will be called "bad." (It will turn out that all of them are full.)

If $(c,B) = 1$ or if $(a,D) = 1$ then the set (3.2) is good. Lemma 2.3 shows that it is full.

If $(c,B) > 1$ and $(a,D) > 1$ the set (3.2) is bad. In order to have a reasonable bound on the upper density of the finite union of these bad sets (3.2), it is necessary to determine how many there are of them.

First compute the number of $B = b + ia$, $0 \le i \le ([a,c]/a) - 1$ that are not relatively prime to $[a,c]$, or, since $(a,b) = 1$, not relatively prime to c. To do so, let $p_1, p_2, \cdots, p_k$ be the distinct primes that divide c but not a. (There may be no such primes.) As i runs through $p_1p_2 \cdots p_k$ consecutive integers, $B = b + ia$ runs through a complete residue system modulo $p_1p_2 \cdots p_k$, of which

$$p_1p_2 \cdots p_k - \phi(p_1p_2 \cdots p_k)$$

are not relatively prime to c. Similarly, let $q_1, q_2, \cdots, q_m$ be the primes that divide a but not c. The number of bad sets of the form (3.2) is then

$$\frac{[a,c]}{a} \cdot \frac{[a,c]}{c}\left(1 - \prod_{p_i}(1 - p_i^{-1})\right)\left(1 - \prod_{q_j}(1 - q_j^{-1})\right) .$$

There may be no bad sets, in which case the proof is already complete.

Each good set of type (3.2) is full, by Lemma 2.3. Each bad set of type (3.2) has upper density at most $1/[a,c]$. Hence, the upper density of the union of the bad sets is at most

$$\frac{1}{[a,c]} \cdot \frac{[a,c]}{a} \cdot \frac{[a,c]}{c}\left(\prod_{p_i \mid ac}(1 - p_i^{-1})\right)^2 ,$$

which equals

$$(1/(a,c))\left(\prod_{p_i \mid ac}(1 - p_i^{-1})\right)^2 .$$

Let

$$k = \left(\prod_{p_i \mid ac}(1 - p_i^{-1})\right)^2 .$$

At the end of the first stage it is known that

$$\underline{\delta} \geq 1/(a,c) - k .$$

Each of the bad sets in stage 1 is then treated as in the example. That is, a bad set

$$\{([a,c]\, x + B)([a,c]\, y + D)\}$$

is written as

$$\left\{(c,B)(a,D)\left(\frac{[a,c]}{(c,B)}\, x + \frac{B}{(c,B)}\right)\left(\frac{[a,c]}{(a,D)}\, y + \frac{D}{(a,D)}\right)\right\}$$

(Keep in mind that $([a,c],B) = (c,B)$ and $([a,c],D) = (a,D)$.)

The set

$$\left\{\left(\frac{[a,c]}{(a,B)}\, x + \frac{B}{(c,B)}\right)\left(\frac{[a,c]}{(a,D)}\, y + \frac{D}{(a,D)}\right)\right\}$$

is then decomposed into products of progressions with equal moduli, as S was. The common modulus is

$$\left[\frac{[a,c]}{(a,B)}, \frac{[a,c]}{(c,D)}\right] , \tag{3.3}$$

whose prime divisors are clearly among the primes that divide $[a,c]$. (In fact, by Lemma 2.1, (3.3) equals $[a,c]$, but this fact plays no role in the proof.)

After completing stage 2, we then have

$$\underline{\delta} \ge 1/(a,c) - k^2 .$$

Similarly, the n^{th} stage shows that

$$\underline{\delta} \ge 1/(a,c) - k^n .$$

Consequently, $\underline{\delta} \ge 1/(a,c)$, and S has density $1/(a,c)$. This concludes the proof.

4. REPRESENTATION OF INTEGERS BY THE FORM $axy + bx + cy$

The theorem has as an immediate consequence the following information about the representation of integers by a certain polynomial form.

Theorem 4.1. Let a be a positive integer, and let b and c be non-negative integers. Then the set of integers expressible in the form

$$axy + bx + cy$$

for some positive integers x and y has a density equal to

$$\frac{a}{(a,b)(a,c)\left(\frac{a}{(a,c)}, \frac{a}{(a,b)}\right)} .$$

Proof. The equation

$$z = axy + bx + cy$$

is equivalent to the equation

$$\begin{aligned} az &= (ax + c)(ay + b) - bc \\ &= (a,c)(a,b)\left(\frac{a}{(a,c)}x + \frac{c}{(a,c)}\right)\left(\frac{a}{(a,b)}y + \frac{b}{(a,b)}\right) - bc . \end{aligned}$$

By the theorem, the set of integers z of the form $axy + bx + cy$ thus has density

$$a \cdot \frac{1}{(a,c)} \cdot \frac{1}{(a,b)} \cdot \frac{1}{\left(\frac{a}{(a,c)}, \frac{a}{(a,b)}\right)} .$$

This completes the proof.

In particular, Theorem 4.1 shows that if $(a,b) = 1$ or if $(a,c) = 1$, then the set $\{axy + bx + cy\}$ has density 1.