AFFINE INVARIANTS, RELATIVELY PRIME SETS, AND A PHI FUNCTION FOR SUBSETS OF $\{1, 2, \ldots, N\}$

Melvyn B. Nathanson[1]
Lehman College (CUNY), Bronx, New York 10468
melvyn.nathanson@lehman.cuny.edu

Received: 8/15/06, Accepted: 12/29/06, Published: 1/3/07

Abstract

A nonempty subset A of $\{1, 2, \ldots, n\}$ is *relatively prime* if $\gcd(A) = 1$. Let $f(n)$ and $f_k(n)$ denote, respectively, the number of relatively prime subsets and the number of relatively prime subsets of cardinality k of $\{1, 2, \ldots, n\}$. Let $\Phi(n)$ and $\Phi_k(n)$ denote, respectively, the number of nonempty subsets and the number of subsets of cardinality k of $\{1, 2, \ldots, n\}$ such that $\gcd(A)$ is relatively prime to n. Exact formulas and asymptotic estimates are obtained for these functions.

Subject class: Primary 11A25, 11B05, 11B13, 11B75.

Keywords: Relatively prime sets, Euler phi function, combinatorial

1. Affine Invariants

Let A be a set of integers, and let x and y be rational numbers. We define the *dilation* $x * A = \{xa : a \in A\}$ and the *translation* $A + y = \{a + y : a \in A\}$. Sets of integers A and B are *affinely equivalent* if there exist rational numbers $x \neq 0$ and y such that $B = x*A+y$. For example, the sets $A = \{2, 8, 11, 20\}$ and $B = \{-4, 10, 17, 38\}$ are affinely equivalent, since $B = (7/3) * A - 26/3$, and A and B are both affinely equivalent to the sets $C = \{0, 2, 3, 6\}$ and $D = \{0, 3, 4, 6\}$. Every set with one element is affinely equivalent to $\{0\}$. Every finite set A of integers with more than one element is affinely equivalent to unique sets C and D of nonnegative integers such that $\min(C) = \min(D) = 0$, $\gcd(C) = \gcd(D) = 1$, and $D = (-1) * C + \max(C)$.

A function $f(A)$ whose domain is the set $\mathcal{F}(\mathbf{Z})$ of nonempty finite sets of integers is called an *affine invariant* of $\mathcal{F}(\mathbf{Z})$ if $f(A) = f(B)$ for all affinely equivalent sets A and B.

[1]This work was supported in part by grants from the NSA Mathematical Sciences Program and the PSC-CUNY Research Award Program.

For example, if $A+A=\{a+a' : a, a' \in A\}$ is the sumset of a finite set A of integers, and if $A-A=\{a-a' : a, a' \in A\}$ is the difference set of the finite set A, then $s(A)=\text{card}(A+A)$ and $d(A)=\text{card}(A-A)$ are affine invariants. More generally, let $u_0, u_1, \ldots, u_n$ be integers and $F(x_1, \ldots, x_n)=u_1x_1+\cdots+u_nx_n+u_0$. Define $F(A)=\{u_1a_1+\cdots+u_na_n+u_0 : a_1, \ldots, a_n \in A \text{ for } i=1, \ldots, n\}$. Then $f(A)=\text{card}(F(A))$ is an affine invariant.

Let $f(A)$ be a function with domain $\mathcal{F}(\mathbf{Z})$. A frequent problem in combinatorial number theory is to determine the distribution of values of the function $f(A)$ for sets A in the interval of integers $\{0, 1, \ldots, n\}$. For example, if $A \subseteq \{0, 1, 2, \ldots, n\}$, then $1 \leq \text{card}(A+A) \leq 2n+1$. For $\ell = 1, \ldots, 2n+1$, we can ask for the number of nonempty sets $A \subseteq \{0, 1, 2, \ldots, n\}$ such that $\text{card}(A+A)=\ell$. Similarly, if $\emptyset \neq A \subseteq \{0, 1, 2, \ldots, n\}$ and $\text{card}(A)=k$, then $2k-1 \leq \text{card}(A+A) \leq k(k+1)/2$, and, for $\ell = 2k-1, \ldots, k(k+1)/2$, we can ask for the number of such sets A with $\text{card}(A+A)=\ell$. In both cases, there is a redundancy in considering sets that are affinely equivalent, and we might want to count only sets that are pairwise affinely inequivalent.

2. Relatively Prime Sets

A nonempty subset A of $\{1, 2, \ldots, n\}$ will be called *relatively prime* if the elements of A are relatively prime, that is, if $\gcd(A)=1$. Let $f(n)$ denote the number of relatively prime subsets of $\{1, 2, \ldots, n\}$. The first 10 values of $f(n)$ are 1, 2, 5, 11, 26, 53, 116, 236, 488, and 983. (This is sequence A085945 in Sloane's *On-Line Encyclopedia of Integer Sequences.*) Let $f_k(n)$ denote the number of relatively prime subsets of $\{1, 2, \ldots, n\}$ of cardinality k. We present exact formulas and asymptotic estimates for $f(n)$ and $f_k(n)$. These estimates imply that almost all finite sets of integers are relatively prime.

No set of even integers is relatively prime. Since there are $2^{[n/2]}-1$ nonempty subsets of $\{2, 4, 6, \ldots, 2[n/2]\}$ and 2^n-1 nonempty subsets of $\{1, 2, \ldots, n\}$, we have the upper bound

$$f(n) \leq 2^n - 2^{[n/2]}. \tag{1}$$

Similarly,

$$f_k(n) \leq \binom{n}{k} - \binom{[n/2]}{k}. \tag{2}$$

If $1 \in A$, then A is relatively prime. Since there are 2^{n-1} sets $A \subseteq \{1, 2, \ldots, n\}$ with $1 \in A$, we have

$$f(n) \geq 2^{n-1}.$$

Let $n \geq 3$. If $1 \notin A$ but $2 \in A$ and $3 \in A$, then A is relatively prime and so

$$f(n) \geq 2^{n-1} + 2^{n-3}.$$

Let $n \geq 5$. If $1 \notin A$ and $3 \notin A$, but $2 \in A$ and $5 \in A$, then A is relatively prime. If $1 \notin A$ and $2 \notin A$, but $3 \in A$ and $5 \in A$, then A is relatively prime. Therefore,

$$f(n) \geq 2^{n-1} + 2^{n-3} + 2 \cdot 2^{n-4} = 2^{n-1} + 2^{n-2}.$$

Similarly,

$$f_k(n) \geq \binom{n-1}{k-1} + \binom{n-3}{k-2} + 2\binom{n-4}{k-2}.$$

3. Exact Formulas and Asymptotic Estimates

Let $[x]$ denote the greatest integer less than or equal to x. If $x \geq 1$ and $n = [x]$, then

$$\left[\frac{x}{d}\right] = \left[\frac{[x]}{d}\right] = \left[\frac{n}{d}\right]$$

for all positive integers d.

Let $F(x)$ be a function defined for $x \geq 1$, and define the function

$$G(x) = \sum_{1 \leq d \leq x} F\left(\frac{x}{d}\right).$$

In the proof of Theorem 1 we use the following version of the Möbius inversion formula (Nathanson [1, Exercise 5 on p. 222]):

$$F(x) = \sum_{1 \leq d \leq x} \mu(d) G\left(\frac{x}{d}\right).$$

Theorem 1 *For all positive integers n,*

$$\sum_{d=1}^{n} f\left(\left[\frac{n}{d}\right]\right) = 2^n - 1 \tag{3}$$

and

$$f(n) = \sum_{d=1}^{n} \mu(d) \left(2^{[n/d]} - 1\right). \tag{4}$$

For all positive integers n and k,

$$\sum_{d=1}^{n} f_k\left(\left[\frac{n}{d}\right]\right) = \binom{n}{k} \tag{5}$$

and

$$f_k(n) = \sum_{d=1}^{n} \mu(d) \binom{[n/d]}{k}. \tag{6}$$

Proof. Let A be a nonempty subset of $\{1, 2, \ldots, n\}$. If $\gcd(A) = d$, then $A' = (1/d) * A = \{a/d : a \in A\}$ is a relatively prime subset of $\{1, 2, \ldots, [n/d]\}$. Conversely, if A' is a relatively prime subset of $\{1, 2, \ldots, [n/d]\}$, then $A = d * A' = \{da' : a' \in A'\}$ is a nonempty subset of $\{1, 2, \ldots, n\}$ with $\gcd(A) = d$. It follows that there are exactly $f([n/d])$ subsets A of $\{1, 2, \ldots, n\}$ with $\gcd(A) = d$, and so

$$\sum_{d=1}^{n} f\left(\left[\frac{n}{d}\right]\right) = 2^n - 1.$$

We apply Möbius inversion to the function $F(x) = f([x])$. For all $x \geq 1$ we define

$$G(x) = \sum_{1 \leq d \leq x} F\left(\frac{x}{d}\right) = \sum_{1 \leq d \leq x} f\left(\left[\frac{x}{d}\right]\right) = \sum_{d=1}^{[x]} f\left(\left[\frac{[x]}{d}\right]\right) = 2^{[x]} - 1$$

and so

$$f([x]) = F(x) = \sum_{1 \leq d \leq x} \mu(d) G\left(\frac{x}{d}\right) = \sum_{d=1}^{[x]} \mu(d) \left(2^{[x/d]} - 1\right).$$

For $n \geq 1$ we have

$$f(n) = \sum_{d=1}^{n} \mu(d) \left(2^{[n/d]} - 1\right).$$

The proofs of (5) and (6) are similar. □

Theorem 2 *For all positive integers n and k,*

$$2^n - 2^{[n/2]} - n2^{[n/3]} \leq f(n) \leq 2^n - 2^{[n/2]}$$

and

$$\binom{n}{k} - \binom{[n/2]}{k} - n\binom{[n/3]}{k} \leq f_k(n) \leq \binom{n}{k} - \binom{[n/2]}{k}.$$

Proof. For $n \geq 2$ we have

$$2^n = f(n) + f([n/2]) + \sum_{d=3}^{n} f\left(\left[\frac{n}{d}\right]\right) + 1 \leq f(n) + 2^{[n/2]} + n2^{[n/3]}.$$

Combining this with (1), we obtain

$$2^n - 2^{[n/2]} - n2^{[n/3]} \leq f(n) \leq 2^n - 2^{[n/2]}.$$

This also holds for $n = 1$.

The inequality for $f_k(n)$ follows similarly from (2) and (5). □

Theorem 2 implies that $f(n) \sim 2^n$ as $n \to \infty$, and so almost all finite sets of integers are relatively prime.

4. A phi Function for Sets

The Euler phi function $\varphi(n)$ counts the number of positive integers $a \leq n$ such that a is relatively prime to n. We define the function $\Phi(n)$ to be the number of nonempty subsets A of $\{1, 2, \ldots, n\}$ such that $\gcd(A)$ is relatively prime to n. For example, for distinct primes p and q we have

$$\Phi(p) = 2^p - 2$$

$$\Phi(p^2) = 2^{p^2} - 2^p$$

and

$$\Phi(pq) = 2^{pq} - 2^q - 2^p + 2.$$

Define the function $\Phi_k(n)$ to be the number of subsets A of $\{1, 2, \ldots, n\}$ such that $\mathrm{card}(A) = k$ and $\gcd(A)$ is relatively prime to n. Note that $\Phi_1(n) = \varphi(n)$ for all $n \geq 1$.

Theorem 3 *For all positive integers n,*

$$\sum_{d|n} \Phi(d) = 2^n - 1. \tag{7}$$

Moreover, $\Phi(1) = 1$ and, for $n \geq 2$,

$$\Phi(n) = \sum_{d|n} \mu(d)\, 2^{n/d} \tag{8}$$

where $\mu(n)$ is the Möbius function. Similarly, for all positive integers n and k,

$$\sum_{d|n} \Phi_k(d) = \binom{n}{k} \tag{9}$$

and

$$\Phi_k(n) = \sum_{d|n} \mu(d) \binom{n/d}{k} \tag{10}$$

Proof. For every divisor d of n, we define the function $\Psi(n, d)$ to be the number of nonempty subsets A of $\{1, 2, \ldots, n\}$ such that the greatest common divisor of $\gcd(A)$ and n is d. Thus,

$$\Psi(n, d) = \mathrm{card}\left(\{A \subseteq \{1, 2, \ldots, n\} : A \neq \emptyset \text{ and } \gcd(A \cup \{n\}) = d\}\right).$$

Then

$$\Psi(n, d) = \Phi\left(\frac{n}{d}\right)$$

and

$$2^n - 1 = \sum_{d|n} \Psi(n, d) = \sum_{d|n} \Phi\left(\frac{n}{d}\right) = \sum_{d|n} \Phi(d).$$

We have $\Phi(1) = 1$. For $n \geq 2$ we apply the usual Möbius inversion and obtain

$$\begin{aligned}\Phi(n) &= \sum_{d|n} \mu(d)\left(2^{n/d} - 1\right) \\ &= \sum_{d|n} \mu(d)\, 2^{n/d} - \sum_{d|n} \mu(d) \\ &= \sum_{d|n} \mu(d)\, 2^{n/d}\end{aligned}$$

since $\sum_{d|n} \mu(n/d) = 0$ for $n \geq 2$.

The proofs of (9) and (10) are similar. $\square$

Theorem 4 *If n is odd, then*

$$\Phi(n) = 2^n + O\left(n2^{n/3}\right)$$

and

$$\Phi_k(n) = \binom{n}{k} + O\left(n\binom{[n/3]}{k}\right).$$

If n is even, then

$$\Phi(n) = 2^n - 2^{n/2} + O\left(n2^{n/3}\right)$$

and

$$\Phi_k(n) = \binom{n}{k} - \binom{n/2}{k} + O\left(n\binom{[n/3]}{k}\right).$$

Proof. We have

$$\begin{aligned}\Phi(n) &= \sum_{\substack{d=1 \\ \gcd(d,n)=1}}^{n} \operatorname{card}\left(\{A \subseteq \{1, 2, \ldots, n\} : A \neq \emptyset \text{ and } \gcd(A) = d\}\right) \\ &= \sum_{\substack{d=1 \\ \gcd(d,n)=1}}^{n} f([n/d]).\end{aligned}$$

Applying Theorem 2, we see that if n is odd, then

$$\begin{aligned}\Phi(n) &= f(n) + f([n/2]) + \sum_{\substack{d=3\\ \gcd(d,n)=1}}^{n} f([n/d]) \\ &= \left(2^n - 2^{[n/2]} + O\left(n2^{n/3}\right)\right) + \left(2^{[n/2]} + O\left(2^{n/4}\right)\right) + O\left(n2^{n/3}\right) \\ &= 2^n + O\left(n2^{n/3}\right).\end{aligned}$$

If n is even, then

$$\begin{aligned}\Phi(n) &= f(n) + \sum_{\substack{d=3\\ \gcd(d,n)=1}}^{n} f([n/d]) \\ &= \left(2^n - 2^{n/2} + O\left(n2^{n/3}\right)\right) + O\left(n2^{n/3}\right) \\ &= 2^n - 2^{n/2} + O\left(n2^{n/3}\right).\end{aligned}$$

These estimates for $\Phi(n)$ also follow from identity (8). The estimates for $\Phi_k(n)$ follow from identity (10). This completes the proof. □

Acknowledgements. I thank Greg Martin for the observation that (4) and (6) follow from (3) and (5) by Möbius inversion, and Kevin O'Bryant for helpful discussions.

References

[1] M. B. Nathanson, *Elementary Methods in Number Theory*, Graduate Texts in Mathematics, vol. 195, Springer-Verlag, New York, 2000.