AN OLD THEOREM ON THE GCD AND ITS APPLICATION TO PRIMES

P. G. Tsangaris
University of Athens, Greece

J. P. Jones*
University of Calgary, Alberta, Canada

(Submitted July 1990)

1. Introduction

We show how an old theorem about the GCD can be used to define primes and to construct formulas for primes. We give new formulas for the characteristic function of the primes, the n^{th} prime p_n, the function $\pi(x)$, and the least prime greater than a given number.

These formulas are all elementary functions in the sense of Grzegorczyk [6] and Kalmar [12] (Kalmar elementary). From a theorem of Jones [11], it will follow that there exist formulas with the same range built up only from the four functions

$$(1.1) \qquad x + y, \qquad [x/y], \qquad x \dot{-} y, \qquad 2^x,$$

by function composition (without sigma signs). There also exist polynomial formulas for the primes, but that is another subject (see [10]).

The constructions here use a theorem of Hacks [7]. (He indicates on page 207 of [7] that this result may have been known to Gauss. See also Dickson [2] vol. 1, p. 333.) Hacks [7] considered sums of the form:

Definition 1.1: $H(k, n) = 2\sum_{i=1}^{n-1}\left\lfloor\frac{k \cdot i}{n}\right\rfloor.$

Here $\lfloor x \rfloor$ denotes the floor (integer part) of x. Hacks proved that sums of this type could be used to define the GCD of k and n, i.e., (k, n).

Theorem 1.1 (Hacks [7]): $H(k, n) = nk - k - n + (k, n)$.

Proof: The proof of this theorem requires the following two lemmas.

Lemma 1.1: Suppose $k \perp n$. Then $H(k, n) = (k - 1)(n - 1)$.

Proof: $ki \equiv kj \pmod{n}$ implies $i \equiv j \pmod{n}$. Thus, the set $\{ki: i = 0, 1, 2, \ldots, n - 1\}$ is a complete residue system mod n. The sum of the remainders in this system must be equal to $n(n - 1)/2$.

Hence, let $ki \equiv r_i \pmod{n}$ where $0 \le r_i < n$, $(i = 1, 2, \ldots, n - 1)$. Then we have

$$ki = \left\lfloor\frac{k \cdot i}{n}\right\rfloor n + r_i \qquad \text{and} \qquad \sum_{i=0}^{n-1} r_i = \frac{n(n - 1)}{2}.$$

Summing the first equation, we find

$$k\sum_{i=0}^{n-1} i = \sum_{i=0}^{n-1} ki = \sum_{i=0}^{n-1}\left\lfloor\frac{k \cdot i}{n}\right\rfloor n + r_i = n\sum_{i=0}^{n-1}\left\lfloor\frac{k \cdot i}{n}\right\rfloor + \sum_{i=0}^{n-1} r_i.$$

Therefore,

$$k\,\frac{n(n - 1)}{2} = \frac{n}{2} \cdot H(k, n) + \frac{n(n - 1)}{2}.$$

*Work supported by Natural Sciences and Engineering Research Council of Canada Research Grant A4525 and the N.S.E.R.C. Program of International Scientific Exchange Awards.

Multiplying by 2 and dividing by n gives the result:

$$k(n - 1) = H(k, n) + n - 1.$$

Lemma 1.2: $H(ad, bd) = adbd - abd + dH(a, b)$.

Proof: Integers i such that $0 \le i \le bd - 1$ can be written in the form $i = bq + j$ where $0 \le q < d$ and $0 \le j < b$. Hence, we have

$$\begin{aligned} H(ad, bd) &= 2\sum_{i=1}^{bd-1}\left\lfloor\frac{ad \cdot i}{bd}\right\rfloor = 2\sum_{i=1}^{bd-1}\left\lfloor\frac{a \cdot i}{b}\right\rfloor = 2\sum_{q=0}^{d-1}\sum_{j=0}^{b-1}\left\lfloor\frac{a(bq+j)}{b}\right\rfloor \\ &= 2\sum_{q=0}^{d-1}\sum_{j=0}^{b-1} aq + 2\sum_{j=0}^{b-1}\sum_{q=0}^{d-1}\left\lfloor\frac{a \cdot j}{b}\right\rfloor = 2\sum_{q=0}^{d-1} abq + 2\sum_{j=0}^{b-1} d\left\lfloor\frac{a \cdot j}{b}\right\rfloor \\ &= 2ab\,\frac{d(d-1)}{2} + 2d \cdot \frac{1}{2}H(a, b) = abd(d-1) + dH(a, b) \\ &= adbd - abd + dH(a, b). \end{aligned}$$

Corollary 1.1 (Hacks [7]): $H(k, n) = nk - k - n + (k, n)$.

Proof: Write $k = ad$ and $n = bd$ where $a \perp b$ and $d = (k, n)$. From Lemma 1.1, we then have $H(a, b) = (a - 1)(b - 1)$. Using this together with Lemma 1.2, we have

$$\begin{aligned} H(k, n) &= H(ad, bd) = adbd - abd + d(a-1)(b-1) \\ &= adbd - abd + abd - da - db + d = nk - k - n + d. \end{aligned}$$

From Corollary 1.1, it follows that the function H is commutative (symmetric), $H(k, n) = H(n, k)$. The function H has other interesting properties. Using an argument similar to the proof of Lemma 1.2, it is easy to show that

$$(1.2) \qquad H(qk, k) = qk(k - 1), \qquad H(k, qk) = qk(k - 1),$$

$$(1.3) \qquad H(qk + r, k) = qk(k - 1) + H(r, k).$$

2. Characteristic Function of the Primes

From Lemmas 1.1 and 1.2, we see that

Lemma 2.1: $1 = (k, n) \Leftrightarrow (k - 1)(n - 1) = H(k, n)$.
$1 < (k, n) \Leftrightarrow (k - 1)(n - 1) < H(k, n)$.

Now let $m = n - 1$ (or $m = \lfloor\sqrt{n}\rfloor$, to be more economical). Then, by Lemma 2.1, n is composite if and only if

$$(\exists k)[1 \le k \le m \text{ and } (k - 1)(n - 1) < H(k, n)].$$

Hence, n is composite if and only if

$$(2.1) \qquad (\exists k)\left[1 \le k \le m \text{ and } 0 < 2\sum_{i=1}^{k-1}\left\lfloor\frac{i \cdot n}{k}\right\rfloor - \sum_{i=1}^{k-1}(n - 1)\right].$$

It follows that n is composite if and only if

$$(2.2) \qquad (\exists k)\left[1 \le k \le m \text{ and } 0 < \sum_{i=1}^{k-1}\left(2\left\lfloor\frac{i \cdot n}{k}\right\rfloor - n + 1\right)\right], \quad m = \lfloor\sqrt{n}\rfloor.$$

When n is prime, these expressions are all 0. So, by summing over k, we can see that n is composite if and only if

$$(2.3) \qquad 0 < \sum_{0 < i < k \le m}\left(2\left\lfloor\frac{i \cdot n}{k}\right\rfloor - n + 1\right), \quad m = \lfloor\sqrt{n}\rfloor.$$

Alternatively, by summing the constant term, (2.3) can be rewritten to say that n is composite if and only if

$$(2.4) \qquad 0 < \sum_{k=1}^{m}\left(\sum_{i=1}^{k-1}\left(2\left\lfloor\frac{i \cdot n}{k}\right\rfloor\right)\right) - \frac{(n-1)(m-1)m}{2}, \quad m = \lfloor\sqrt{n}\rfloor.$$

This is equivalent to the statement that n is composite if and only if

$$(2.5) \qquad 0 < m(m-1)(1-n) + \sum_{0<i<k\le m} 4\left\lfloor\frac{i \cdot n}{k}\right\rfloor, \quad m = \lfloor\sqrt{n}\rfloor.$$

Since these expressions are zero when n is a prime, they characterize primes. We summarize (2.3) in Theorem 2.1.

Theorem 2.1: Let $g(n)$ be defined by

$$(2.6) \qquad g(n) = \sum_{0<i<k\le m}\left(2\left\lfloor\frac{i \cdot n}{k}\right\rfloor - n + 1\right),$$

where $m = \lfloor\sqrt{n}\rfloor$ or $m = n - 1$. Then, for all $n > 1$, n is prime if and only if $g(n) = 0$. And n is composite if and only if $g(n) \ge 1$.

The subtraction function $x \mathbin{\dot{-}} y$ or the sgn(x) function can now be used to obtain a characteristic function for the primes. A *characteristic function* for a set is a two-valued function taking value 1 on the set and value 0 on the complement of the set.

The proper subtraction function $x \mathbin{\dot{-}} y$ is defined to be $x - y$ for $y \le x$ and 0 for $x < y$. The sign function sgn(x) is defined by sgn(x) = +1 if $x > 0$, by sgn(x) = −1 if $x < 0$ and sgn(0) = 0.

Now define $h(n)$ to be $h(n) = 1 \mathbin{\dot{-}} g(n)$ or define $h(n) = 1 - \operatorname{sgn} g(n)$. Then it follows from Theorem 2.1 that $h(n)$ is a characteristic function for the set of primes.

Theorem 2.2: Let $h(n)$ be defined by

$$(2.7) \qquad h(n) = 1 \mathbin{\dot{-}} \sum_{0<i<k\le m}\left(2\left\lfloor\frac{i \cdot n}{k}\right\rfloor - n + 1\right),$$

where $m = n - 1$ or $m = \lfloor\sqrt{n}\rfloor$. Then n is prime if and only if $h(n) = 1$. And n is composite if and only if $h(n) = 0$. (These statements hold for $n > 1$.)

We can use the function h to construct a formula for the function $\pi(x)$, [$\pi(x)$ = the number of primes $\le x$]. From Theorem 2.2, we have

Theorem 2.3: The function $\pi(x)$ is given by

$$(2.8) \qquad \pi(x) = \sum_{n=2}^{x} h(n) = \sum_{n=2}^{x}\left(1 \mathbin{\dot{-}} \sum_{0<i<k\le m}\left(2\left\lfloor\frac{i \cdot n}{k}\right\rfloor - n + 1\right)\right).$$

Proof: The idea of (2.8) is that the characteristic function h counts the primes $\le x$. [We start the sum at $n = 2$ instead of at $n = 1$ because $h(1) = 1$.]

3. Formula for the n^{th} Prime

Define $C(a, n) = 1 \mathbin{\dot{-}} (a + 1) \mathbin{\dot{-}} n$. Then $C(a, n)$ is the characteristic function of the relation $a < n$. That is, if $a < n$, then $C(a, n) = 1$. If $n \le a$, then $C(a, n) = 0$.

Now $\pi(i) < n$ if and only if $i < p_n$. Hence $C(\pi(i), n) = 1$ if and only if $i < p_n$. The n^{th} prime p_n is therefore given by the following formula:

$$(3.1) \qquad p_n = \sum_{\ell=0}^{k} C(\pi(\ell), n) = \sum_{\ell=0}^{k}\left(1 \mathbin{\dot{-}} (\pi(\ell) + 1 \mathbin{\dot{-}} n)\right)$$

when k is large enough ($k \geq p_n - 1$). It is known that

$$p_n < n(\log(n) + \log(\log(n))) \text{ for } n > 5$$

(see Rosser & Schoenfeld [13]). So we can take

$$k = n^2 \quad \text{or} \quad k = 2n \log(n + 1).$$

Using Theorem 2.3 and the fact that $h(1) = 1$, we have

$$(3.2) \qquad p_n = \sum_{\ell=0}^{k}\left(1 \dot{-} \left(\left(\sum_{j=1}^{\ell} h(j)\right) \dot{-} n\right)\right).$$

Thus we have, from Theorem 2.2,

Theorem 3.1: The n^{th} prime, p_n, is given by

$$(3.3) \qquad p_n = \sum_{\ell=0}^{k}\left(1 \dot{-} \left(\left(\sum_{j=1}^{\ell}\left(1 \dot{-} \sum_{0<i<k\leq m}\left(2\left\lfloor\frac{i \bullet j}{k}\right\rfloor - j + 1\right)\right)\right) \dot{-} n\right)\right), \text{ for } n > 1.$$

4. Next Prime Greater than a Given Number

The function g of Theorem 2.1 has the property that it is nonnegative and $g(n) = 0$ if and only if n is a prime. The function h also has this property. Hence, we can use either h or g in the following construction of a formula for the next prime greater than a number q. (The number q can be any integer, it does not need to be a prime.)

Theorem 4.1: The next prime greater than q is given by the function

$$(4.1) \qquad N(q) = \sum_{j=0}^{2q}\left(1 \dot{-} \sum_{n=0}^{j}(n \dot{-} q)(1 \dot{-} g(n))\right).$$

Proof: From Bertrand's Postulate, we know that for every $q \geq 1$ there is a prime p such that $q < p \leq 2q$. Fix q and let p denote the least such prime p. Put

$$f(n) = (n \dot{-} q)(1 \dot{-} g(n)).$$

Then $f(n) \geq 0$ for $n \geq 0$. Also $f(n) > 0$ if and only if $q < n$ and $g(n) = 0$, i.e., if and only if n is a prime greater than q. Now

$$1 \dot{-} (f(0) + f(1) + \cdots + f(j)) = 1 \text{ for } j < p.$$

But

$$1 \dot{-} (f(0) + f(1) + \cdots + f(j)) = 0 \text{ for } p < j.$$

Hence, p is equal to the sum

$$(4.2) \qquad N(q) = \sum_{j=0}^{2q}\left(1 \dot{-} (f(0) + f(1) + \cdots + f(j))\right),$$

a sum of exactly p ones.

Bertrand's Postulate is a theorem that was proved by P. L. Chebychev in the nineteenth century. See Hardy [8, p. 349] for a modern proof (due to Erdös [4]).

References

1. J. Bertrand. *Jour. de l'ecole Roy. Polyt.* cah. 30, tome 17 (1845), p. 129.
2. L. E. Dickson. *History of the Theory of Numbers.* New York: Carnegie Institute of Washington, 1919, 1920, 1923.
3. U. Dudley. "History of a Formula for Primes." *Amer. Math. Monthly* 76 (1969): 23-28.

4. P. Erdös. *Acta Litt. Ac. Sci.* (Szeged) 5 (1932):194-98.
5. R. L. Goodstein & C. P. Wormell. "Formulae for Primes." *The Math. Gazette 51* (1967):35-38.
6. A. Grzegorczyk. "Some Classes of Recursive Functions." *Rozprawy Mathematyczne* no. 4, Institut Matematyczny Polskiej Akademie Nauk, Warszawa. MR 15 66.
7. J. Hacks. "Über einige für Primzahlen charakteristische Beziehungen." *Acta Mathematica 17* (1893):205-08 (Mittag-Leffler, Stockholm, Uppsala).
8. G. H. Hardy. *An Introduction to the Theory of Numbers*. London: Oxford University Press, 1960.
9. J. P. Jones. "Formula for the n^{th} Prime Number." *Canadian Math. Bull. 18* (1975):433-34.
10. J. P. Jones, D. Sato, H. Wada, & D. Wiens. "Diophantine Representation of the Set of Prime Numbers." *Amer. Math. Monthly 83* (1976):449-64. MR 54 2615.
11. J. P. Jones. "Basis for the Kalmar Elementary Functions." *Number Theory and Applications*, ed. R. A. Mollin. NATO Advanced Study Institute, April 27-May 5, 1988, Banff Alberta. NATO ASI Series. Dordrecht, Netherlands: Kluwer Publishing Co., 1989, pp. 435-44.
12. L. Kalmar. "Egyszerü pelda eldönthetetlen arithmetikai problemara" (Hungarian with German abstract). *Mate es Fizikai Lapok 50* (1943):1-23.
13. J. Barkley Rosser & Lowell Schoenfeld. "Approximate Formulas for Some Functions of Prime Numbers." *Illinois J. Math. 6* (1962):64-94.
14. P. G. Tsangaris. "Prime Numbers and Cyclotomy-Primes of the Form $x^2+(x+1)^2$." Ph.D. Thesis, Athens University, Athens, 1984 (in Greek, English summary).
15. C. P. Willans. "On Formulae for the n^{th} Prime Number." *The Math. Gazette 48* (1964):413-15.

AMS Classification numbers: 11A51, 11Y11.

*****