NOTES

Formulas for Primes

UNDERWOOD DUDLEY
DePauw University
Greencastle, IN 46135

Formulas fascinate. Not everyone, of course, and not even every mathematician, but for some an array of symbols containing an equals sign has more allure and power than the same thing expressed in any other way. The reason for this probably sprouts in one's mathematical infancy, when we discover that manipulating symbols and using formulas can actually give us answers to problems. To the immature mathematical mind it seems almost magical; it seems as if we are getting something for nothing! First loves are not forgotten, so it is no surprise that a love of formulas, conscious or not, should persist into a person's mathematical adulthood.

Primes fascinate. Not everyone, of course, and not even every mathematician, but for some, contemplating their irregular but tantalizing not-quite-random march through the integers can lead to mystical experiences, or at least a desire to find some order in their chaos. Some victims of both of these fascinations look for formulas: formulas for p_n, **the nth prime**, or formulas for $\pi(n)$, **the number of primes less than or equal to n**, or formulas which give prime values exclusively. It is the purpose of this note to survey formulas for primes and show their wide variation from worthless, to interesting, to astonishing. I do not claim to include every formula ever found; in fact, a complete catalog would have little point. In weighing this representative sample, it seems that the balance tilts more to the side of worthless than astonishing. Although fascinating, a search for formulas for primes should be viewed in the same way as trying to find an elementary proof of Fermat's Last Theorem: good for recreation but almost certain to be fruitless.

Formulas for primes are not new, nor are they very old. Euler's curious polynomial $n^2 - n + 41$, which produces a prime for every integer n from 1 to 40, dates back to 1772. Not long after, Legendre and Gauss were trying to estimate $\pi(n)$, but it was not until the 1890s that authors started to publish formulas. They have, however, continued ever since. The first publications give formulas for χ, the **characteristic function for primes**, which is defined:

$$\chi(n) = \begin{cases} 1 & \text{if } n \text{ is prime} \\ 0 & \text{if } n \text{ is composite.} \end{cases}$$

An 1895 example is

$$\chi(n) = \frac{e^{2\pi i (n-1)!/n} - 1}{e^{-2\pi i/n} - 1}$$

[**D**, p. 427]. (For several reasons, some of our references are to secondary sources, hence we will use a slightly unorthodox method of noting these. References to volume I of Dickson's *History* [**1**] are noted by a **D** and those referring to items appearing in *Mathematical Reviews* are noted by an **MR**.)

Wilson's Theorem says that *n is prime if and only if* $(n-1)! \equiv -1$ *mod* n. This fact is the backbone of many formulas for functions which are cousin to $\chi(n)$. These functions have the property that they take on one type of value when n is prime and another type of value otherwise. A 1911 example [**D**, p. 428] is

$$f(n) = \sin^2 \pi n + \sin^2 \pi \left(\frac{1 + (n-1)!}{n} \right), n > 1; \tag{1}$$

$f(n)$ is zero if and only if n is prime. Another formula for the function $f(n)$ which does not depend on Wilson's Theorem is [**D**, p. 428]

$$f(n) = \frac{\sin^2 \pi n}{(\pi n)^2 (1 - n^2)^2} \cdot \sum_{k=2}^{\infty} \frac{\pi n}{k \sin \pi n / k}, n > 1.$$

Although Wilson's Theorem has proved to be notoriously unuseful in finding primes, this has not stopped the production of formulas in which it is essential. In a recent example, an author [**MR** 31(1966) #2198] takes the function

$$g(n) = \cos^2 \left(\pi \frac{(n-1)! + 1}{n} \right)$$

($g(n)$ is an integer if and only if n is prime) and uses this formula to produce a formula for p_{n+1}. Some other variations for the function $g(n)$ are [**16**],

$$g_1(n) = \frac{\sin^2(\pi (n-1)!/n)}{\sin^2(\pi/n)};$$

also [**MR** 45(1973) #8601]

$$g_2(n) = \frac{1 - \cos(\pi (n-1)!/n)}{1 + \cos(\pi/n)}$$

and [**MR** 21(1960) #7179]

$$g_3(n) = \frac{(j-1)!(n-j)! + (-1)^{j+1}}{n}.$$

Other ideas can be used too. Fermat's Theorem, which says that $a^{p-1} \equiv 1 \bmod p$ if p is prime and $(a, p) = 1$, is the idea in the formula [**MR** 40(1970) #4197]

$$\chi(n) = \prod_{2 \leqslant p \leqslant \sqrt{n}} \frac{1 - \cos(2 n^{p-1} \pi / p)}{1 - \cos(2\pi/p)}$$

where $n \geqslant 4$ and the product is taken over primes. If n is composite, one of the primes will divide it, making the product zero; if n is prime, all of the factors in the product are 1. Another idea is to observe that $\lfloor n/i \rfloor < n/i$ for $i = 2, 3, \ldots, n-1$ if and only if n is prime (here $\lfloor x \rfloor$ denotes the greatest integer not greater than x). An author [**MR** 51(1976) #12676] used this to produce the formula

$$\chi(n) = \left\lfloor \left(\sum_{i=1}^{n-1} \lfloor \lfloor n/i \rfloor / (n/i) \rfloor \right)^{-1} \right\rfloor.$$

Once you have something like these, it is easy to play the formula game. For example, two authors, seventeen years apart [**D**, p. 432], [**D**, p. 434] noted that if in (1) you replace n by x, replace $(n-1)!$ by $\Gamma(x)$ and call the expression $f(x)$, then

$$\pi(n) = \frac{1}{2\pi i} \int_C \frac{f'(x)}{f(x)} dx$$

where C is a closed contour including the x-axis from 1 to n and excluding any complex zeros of f: a formula, but not one which has been of any computational use. As has been noted [**16**], once you have $\chi(n)$, you immediately have a formula for $\pi(n)$:

$$\pi(n) = \sum_{k=2}^{n} \chi(k). \tag{2}$$

If you also utilize C_n, the **characteristic function for the interval** $[0, n]$, that is, $C_n(k) = 1$ for $k \leqslant n$ and 0 for $k > n$, then you have a formula for p_{n+1}:

$$p_{n+1} = 2 + \sum_{k=2}^{\infty} C_n(\pi(k)). \tag{3}$$

The function C_n can be given an impressive analytic representation as an integral, or can be expressed by the combinatorial formula

$$C_n(k) = 1 + \sum_{j=0}^{k-n-1} (-1)^{j+1} \binom{n+j}{j} \binom{k}{n+1+j} \tag{4}$$

(where $\binom{s}{t} = 0$ if $t > s$). The function $\phi(n)$, which counts the number of positive integers less than n and relatively prime to n (called Euler's ϕ-function), and the floor function $\lfloor x \rfloor$ can be combined to give an equally impressive formula for $\pi(k)$:

$$\pi(k) = \sum_{i=2}^{k} \left\lfloor \frac{\phi(i)}{i-1} \right\rfloor. \tag{5}$$

The simple fact that $\phi(i) = i - 1$ if and only if i is prime makes (5) transparent. Now you can substitute (4) and (5) into (3) and there you are with a brand-new formula for p_{n+1}. Or you may take your favorite expression for $\chi(n)$, substitute it into (2) and obtain a new formula for $\pi(n)$. You may use these concoctions as your own, with no charge, to impress those who are impressed by such things.

The first formula for p_{n+1} appeared in 1900 [7]. It was based on the expression

$$F(n,k) = \frac{k!}{P(n,k)} + \frac{P(n,k)}{(k-1)!} - \left\lfloor \frac{(k-1)!}{P(n,k)} \right\rfloor \tag{6}$$

where

$$P(n,k) = 2^{e_1} 3^{e_2} 5^{e_3} \cdots p_n^{e_n}$$

where

$$e_i = \lfloor k/p_i \rfloor + \lfloor k/p_i^2 \rfloor + \lfloor k/p_i^3 \rfloor + \cdots .$$

For $2 \leqslant k < p_{n+1}$ it is well known that $P(n,k) = k!$, so (6) says that $F(n,k) = 1 + k - 0 = k + 1$. But when $k = p_{n+1}$, $P(n,k) = (k-1)!$ and so (6) says that $F(n,k) = k + 1 - 1 = k$. From these observations, with a little work, it is possible to get a formula which starts "$p_{n+1} = \cdots$." But the formula is a restatement of the fact that the prime-power decomposition of $k!$ differs from that of $(k-1)!$ by exactly one factor if and only if k is prime, and is just as useful for finding primes. The idea was not new, for another author [**D**, p. 437] had noted the previous year that p_{n+1} is the only solution greater than 1 of the equation

$$x! = x \prod_{i=1}^{n} p_i^{\lfloor x/p_i \rfloor + \lfloor x/p_i^2 \rfloor + \cdots}$$

I am not sure if G. H. Hardy was being satirical when he gave [**D**, p. 438] a formula for the largest prime dividing a positive integer x:

$$\lim_{r \to \infty} \lim_{m \to \infty} \lim_{n \to \infty} \sum_{t=0}^{m} \left(1 - \left(\cos\left[(t!)^r \pi/x\right]\right)^{2n}\right).$$

It takes more than a little effort to understand the computationally useless formula, but none to understand the words. In any event, the joke, if a joke it was, did not kill the genre. The author of a recent example [**MR** 52(1976) #246, where the formula is misquoted] defines (using Wilson's Theorem again)

$$f(n) = \operatorname{sgn}\left(\frac{2(n-1)!}{n} - \left\lfloor \frac{2(n-1)!}{n} \right\rfloor\right);$$

here $\operatorname{sgn}(x) = -1$, 0, or 1 for x negative, zero, or positive respectively. This is an opaque way of writing "$f(n) = 1$ if n is an odd prime and 0 otherwise." He then writes that for $n \geqslant 2$

$$\begin{aligned} p_{n+1} = {} & (p_n+2)f(p_n+2) + (p_n+4)f(p_n+4)(1-f(p_n+2)) \\ & + (p_n+6)f(p_n+6)(1-f(p_n+2))(1-f(p_n+4)) \\ & + (p_n+8)f(p_n+8)(1-f(p_n+2))(1-f(p_n+4))(1-f(p_n+6)) \\ & + \cdots . \end{aligned}$$

Notice that all of the terms except one on the right are zero. Since there is no chance of evaluating $f(n)$ other than by knowing when n is prime, the formula amounts to no more than the statement that p_{n+1} is the first prime after p_n. As H. S. Wilf has recently observed in a similar context [**15**], this latter statement is unacceptable in polite society. That the formula appeared in a respected journal is testimony to the power of formulas. A similar result appeared in 1950 [**MR** 12(1951) #392]:

$$\pi(n) = n - \frac{1}{2\pi}\int_0^{2\pi} \sum_{m=1}^{n} \cos(xP(m))\,dx \tag{7}$$

where

$$P(m) = \prod_{j=1}^{m-1}\prod_{k=1}^{m-1}(jk-m)$$

and $P(1) = 0$. For $m \geqslant 2$, $P(m) = 0$ exactly when m is composite, so the second term on the right of (7) is just the number of composite integers not greater than n. Unless some other way is found to compute $P(m)$ (very doubtful) the formula is a truism. Another author (cited in [**14**, p. 149]) gave a formula somewhat reminiscent of Hardy's: for $n \geqslant 3$,

$$\pi(n) = 1 + \sum_{k=3}^{n}\left(1 - \lim_{m\to\infty}\left(1 - \prod_{j=2}^{k-1}\sin^2(k\pi/j)\right)^m\right).$$

I leave to the reader the task of deciphering why the formulas work, and the task of computing with them.

Here is the other extreme. W. H. Mills [**11**] proved that *there is a real number A such that* $\lfloor A^{3^n} \rfloor$ *is prime for* **all** n, $n = 1, 2, 3, \ldots$. Isn't that astonishing? Doesn't it make you wonder how it could be true? The proof shows that we do not have an infinite prime-generator because we cannot find A by any means other than constructing it, and to construct it we need to be able to recognize arbitrarily large primes. The result was so striking that it provoked a large number of papers [**MR** 11(1950) #664], [**MR** 13(1952) #321], [**MR** 13(1952) #321a], [**MR** 14(1953) #256], [**MR** 15(1954) #11], including a proof by E. M. Wright [**17**] that there are infinitely many suitable numbers A and moreover, the set of all of them has cardinality c, measure 0, and is nowhere dense. A satisfying result; we now know the odds of finding A. A summary of the above activity can be found in [**2**].

The following idea was discovered independently at about the same time by two authors [**MR** 14(1953) #355], [**MR** 14(1953) #621]. Let

$$s = 0.2003000050000007000000110\ldots = \sum_{n=1}^{\infty} p_n/10^{n^2}.$$

The real number s contains all the primes separated by many zeros, and the formula

$$p_n = \lfloor 10^{n^2}s \rfloor - 10^{2n-1}\lfloor 10^{(n-1)^2}s \rfloor$$

retrieves them. Neither author was aware that Leo Moser had done the same thing earlier [**12**] using the number $\Sigma_{n=1}^{\infty} p_n/10^{n(n+1)/2}$ instead, calling his own work "admittedly rather trivial." Trivial or not, both results were later generalized [**MR** 24(1962) #A1869]. The idea of putting the primes into a real number so as to get them out later could have come from Mills' theorem.

I suspect that the late J. M. Gandhi thought that his formula for primes [3] actually had some chance of being useful. He proved that if $P_n = p_1 p_2 \cdots p_n$, then p_{n+1} is the unique integer m satisfying the inequality

$$1 < 2^m \left(\sum_{d \mid P_n} \mu(d)/(2^d - 1) - 1/2 \right) < 2 \tag{8}$$

(here μ is the Mobius function: $\mu(1) = 1$ and for $d > 1$, $\mu(d) = 0$ if d has a square factor; if not, $\mu(d) = 1$ if d has an even number of prime divisors and $\mu(d) = -1$ otherwise). The sum in (8) is finite, it involves no primes larger than p_n, and perhaps some way of evaluating it easily could be found. But S. W. Golomb later showed [4] that the formula is a version of the Sieve of Eratosthenes. Another hope dashed! Golomb [5] used similar ideas to get other formulas, such as

$$p_{n+1} = \lim_{s \to 0} \left(P_n(s) \zeta(s) - 1 \right)^{-1/s}$$

where

$$P_n(s) = \prod_{i=1}^{n} (1 - p_i^{-s}) \text{ and } \zeta(s) = \sum_{n=1}^{\infty} n^{-s}.$$

Other formulas based on the Sieve had appeared earlier ([**MR** 15(1954) #685], [**MR** 26(1963) #1289], and [**MR** 27(1964) #101] as was noted in [13]). They are still coming [**MR** 81j #10008].

Though not a formula, the following result of Mann and Shanks [10] is remarkable. Write Pascal's triangle with row k starting in column $2k$:

		Column														
		0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
	0	1														
	1		1	1												
	2					1	2	1								
	3							1	3	3	1					
Row	4									1	4	6	4	1		
	5											1	5	10	10	5
	6													1	6	15
	7															1 ....

They proved that in this array: *a column number is prime if and only if each number in it is divisible by the corresponding row number.* Is that not charming? H. W. Gould later showed [6] that the result is equivalent to the statement: n is prime if and only if k divides $\binom{k}{n-2k}$ for all $k > 1$ such that $n/3 \leqslant k \leqslant n/2$. Another pleasing result [9] is a suggested algorithm to find the next prime after p: add to p the smallest positive integer i not of the form $i \lfloor (p+i)/i \rfloor - p$ for $i = 2, 3, \ldots, p$. This works assuming the truth of the almost surely true but unproven statement that there is a prime between n and $n + n^{1/2}$ for n sufficiently large.

Deep results on Diophantine sets resulting from work on Hilbert's Tenth Problem have made it possible to prove that the set of primes is exactly the set of positive values of some polynomial. In [8] such a polynomial, of degree 25 in 26 variables, is actually written out. That too is startling; you might wonder why no computer has been set to work substituting numbers in it, since every time a value is positive it must be prime. The reason is that the form of the polynomial is

$$(x_{11} + 2)\left(1 - \sum_{i=1}^{14} \left(P_i(x_1, x_2, \ldots, x_{26}) \right)^2 \right)$$

so it is positive only if fourteen rather complicated polynomials simultaneously vanish. Hence the first positive value might not appear until considerably after the end of the universe, and even then it might be something trivial, like 17.

The conclusion to be drawn from all this, I think, is that formulas for formulas' sake do not advance the mathematical enterprise. Formulas should be useful. If not, they should be astounding, elegant, enlightening, simple, or have some other redeeming value. (Conway's prime-producing machine, this MAGAZINE, pp. 26-33, succeeds on all counts.) Authors who discover formulas should not rush into print. Even as in business and marriage, in mathematics not *all* that is true needs to be published. Gauss, as always, had it right: *pauca sed matura*.

The author wishes to express awe at the magnitude of the referee's efforts and gratitude for the many helpful suggestions.

References

[1] L. E. Dickson, History of the Theory of Numbers, Chelsea, New York, 1952; reprint of the 1919 edition.
[2] U. Dudley, History of a formula for primes, Amer. Math. Monthly, 76(1969) 23–28. MR 38(1969) #4270.
[3] J. M. Gandhi, Formulae for the *n*th prime, Proc. Washington State Univ. Conf. on Number Theory, Washington State Univ., 1971, pp. 96–106. MR 48(1974) #218.
[4] S. W. Golomb, A direct interpretation of Gandhi's formula, Amer. Math. Monthly, 81(1974) 752–754. MR 50(1975) #7003.
[5] ______, Formulas for the next prime, Pacific J. Math., 63(1976) 401–404. MR 53(1977) #13094.
[6] H. W. Gould, A new primality criterion of Mann and Shanks, Fibonacci Quart., 10(1972) 355–364, 372. MR 47(1974) #119.
[7] C. Isenkrahe, Ueber eine Losung der Aufgabe, jede Primzahl als function der vorhergehenden Primzahlen durch einen geschlossenen Ausdruck darzustellen, Math. Ann., 53(1900) 42–44.
[8] J. P. Jones, D. Sato, H. Wada, D. Wiens, Diophantine representation of the set of prime numbers, Amer. Math. Monthly, 83(1976) 449–464. MR 54(1977) #2615.
[9] Steven Kahan, On the smallest prime greater than a given positive integer, this MAGAZINE, 47(1974) 91–93. MR 48(1974) #10964.
[10] H. B. Mann and Daniel Shanks, A necessary and sufficient condition for primality and its source, J. Combin. Theory Ser. A., 13(1972) 131–134. MR 46(1973) #5225.
[11] W. H. Mills, A prime representing function, Bull. Amer. Math. Soc., 53(1947) 604. MR 8(1947) #567.
[12] Leo Moser, A prime representing function, this MAGAZINE, 23(1950) 163–164.
[13] K. S. S. Namboordiripad, A note on formulae for the *n*th prime, Monatsh. Math., 75(1971) 256–262. MR 46(1973) #126.
[14] Joe Roberts, Elementary Number Theory, MIT Press, Cambridge, 1977. MR 58(1979) #16472.
[15] H. S. Wilf, What is an answer?, Amer. Math. Monthly, 89(1982) 289–292.
[16] C. P. Willans, On formulae for the *n*th prime number, Math. Gaz., 48(1964) 413–415. MR 31(1966) #2198.
[17] E. M. Wright, A class of representing functions, J. London Math. Soc., 29(1954) 63–71. MR 15(1954) #288.

***Prime* observation 1. The forest *prime*val.**