MATHEMATICS OF COMPUTATION
Volume 77, Number 261, January 2008, Pages 495–502
S 0025-5718(07)02002-9
Article electronically published on May 14, 2007

THE LUCAS-PRATT PRIMALITY TREE

JONATHAN BAYLESS

Abstract. In 1876, E. Lucas showed that a quick proof of primality for a prime p could be attained through the prime factorization of $p-1$ and a primitive root for p. V. Pratt's proof that PRIMES is in NP, done via Lucas's theorem, showed that a certificate of primality for a prime p could be obtained in $O(\log^2 p)$ modular multiplications with integers at most p. We show that for all constants $C \in \mathbb{R}$, the number of modular multiplications necessary to obtain this certificate is greater than $C \log p$ for a set of primes p with relative asymptotic density 1.

1. Introduction

Over a quarter of a century before Agrawal, Kayal, and Saxena showed PRIMES is in P in [1], V. Pratt showed in [11] that PRIMES is in NP by utilizing the following theorem of Lucas:

Theorem 1. *Suppose $p > 1$ is an odd integer and*

$$\begin{cases} a^{\frac{p-1}{2}} \equiv -1 \mod p, \\ a^{\frac{p-1}{2q}} \not\equiv -1 \mod p \text{ for every odd prime } q \mid p-1. \end{cases} \tag{1.1}$$

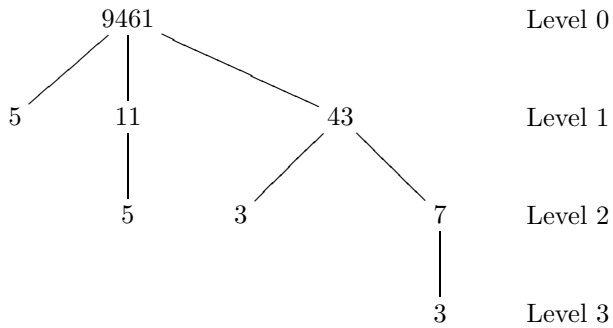
Then p is prime and a is a primitive root of p. Conversely, if p is an odd prime, then every primitive root a of p satisfies conditions (1.1).

In 1877, Lucas [8] stated a result essentially equivalent to Theorem 1; it is based on his work from a year earlier. The actual statement presented here can be found in [3, Theorem 4.1.8].

To achieve this Lucas-Pratt certificate of primality, we would need to find a primitive root a modulo p, and we would need to certify each of the primes q appearing in the second condition. By charting this process, we find what is called a Lucas-Pratt tree. For example, to certify 9461 in this manner, we have the following tree:

Received by the editor June 26, 2006 and, in revised form, November 14, 2006.
2000 *Mathematics Subject Classification.* Primary 11Y16; Secondary 11N37.
The author was supported by a Dartmouth Graduate Fellowship.

©2007 American Mathematical Society
Reverts to public domain 28 years from publication



Notice that the branches beneath a prime r correspond to the odd primes dividing $\varphi(r) = r-1$. Pratt showed that the number of primes appearing in the Lucas-Pratt tree for a prime p, denoted $N(p)$, is $O(\log p)$. In fact, he showed:

Theorem 2. $N(p) < \log_2 p$.

Here $\log_2 p$ denotes the base-2 logarithm of p.

Let $M(p)$ denote the number of modular multiplications of integers with size at most p (using a standard modular exponentiation algorithm) needed to verify the primality of an integer p by this method. As a corollary to the above theorem, Pratt showed:

Corollary 1. $M(p) = O(\log^2 p)$.

See [11] for details of either proof, or [3, Theorem 4.1.9] for a simple exposition of both results. It is clear that $M(p)$ is at least $\log_2 p - 2$, as this is a lower bound for the number of multiplications necessary to evaluate the first condition of Theorem 1.

A relevant result of Pomerance is the following theorem (found in [10]):

Theorem 3. *For every prime p there is a proof that it is prime which requires for its verification $(\frac{5}{2} + o(1)) \log_2 p$ multiplications modulo p.*

Thus, in principle, for each prime p there exists a primality proof of order $O(\log p)$, and so one might ask if Pratt's result can be improved to $M(p) = O(\log p)$.

For any positive integer n, let $\varphi_k(n)$ denote the k-th iterate of the Euler φ-function at n, and let $\varphi_0(n) = n$. Define $F(p)$ as the product of the distinct odd primes dividing $\prod_{k\geq 0} \varphi_k(p)$. An odd prime $q \neq p$ appears in level $k > 1$ of the Lucas-Pratt tree for p if it divides $\varphi(r)$ for some prime r in level $k-1$ of the tree. Since this means that $q \mid \varphi_k(p)$, we have that $F(p)$ is equal to the squarefree product of the primes appearing in the Lucas-Pratt tree for p. As an example, $F(9461) = 9461 \cdot 5 \cdot 11 \cdot 43 \cdot 3 \cdot 7 = 469880565$.

We first prove a theorem analogous to a result in [7] and suggested in the same paper. The proof of this theorem follows the same plan as in [7], but certain extra tools are brought into play.

Theorem 4. *For each $\epsilon > 0$, the set of primes p for which*

$$\varphi(F(p)) > p^{(1-\epsilon)\frac{\log\log p}{\log\log\log p}}$$

has relative asymptotic density 1 within the set of all primes.

We use this theorem to obtain the following result:

Theorem 5. *Let $\epsilon > 0$ be arbitrary and let $C(p) = (1-\epsilon)\log\log p/\log\log\log p$. Then $M(p) > C(p)\log_2 p$ for a set of primes p with relative asymptotic density 1. In particular, for all $C \in \mathbb{R}$, $M(p) > C\log p$ for almost all primes p.*

As noted in Pratt's paper, naive multiplication modulo p has bit complexity $O(\log^2 p)$, and so, using naive multiplication, the bit complexity of the certificate of the prime p is $O(\log^4 p)$. One might ask if the bit complexity can be bounded below in a manner analogous to what we do here with $M(p)$. Limited attempts to use the result of this paper to answer this question have met with little success, but it would be interesting to see if a similar lower bound on the bit complexity is possible.

In [2], the related question of the height of the Lucas-Pratt tree is considered by Banks and Shparlinski, who prove a lower bound on the height for almost all primes p.

Throughout, p, q, and r will always be primes, and n will always be a positive integer.

2. Technical preparation

We now begin developing the tools we will use to prove Theorem 4.

Lemma 1. *For any value of $x \geq e^e$ and any prime $p < x$, the following holds:*

$$\sum_{\substack{r \leq \frac{x}{2} \\ r \equiv 1 \bmod p}} \frac{1}{r\log\frac{x}{r}} \leq \frac{30\log\log x}{p\log\frac{x}{p}}.$$

Proof. We use the Brun-Titchmarsh inequality of [9]; namely, for coprime integers k and ℓ, the number of primes $q \leq x$ with $q \equiv \ell \bmod k$, denoted $\pi(x;k,\ell)$, satisfies

$$\pi(x;k,\ell) \leq \frac{2x}{\varphi(k)\log\frac{x}{k}} \quad \text{for} \quad x > k.$$

For prime $p \geq 2$ and $(\ell, p) = 1$, we can use the fact that $\varphi(p) = p-1$ to see that

$$\pi(x;p,\ell) \leq \frac{2x}{\varphi(p)\log\frac{x}{p}} \leq \frac{4x}{p\log\frac{x}{p}} \quad \text{for} \quad x > p. \tag{2.1}$$

Note that the first prime $r \equiv 1 \bmod p$ is at least $\frac{3p}{2}$, so we need only consider $p \leq \frac{x}{3}$. We use partial summation to obtain

$$\begin{aligned}
\sum_{\substack{r \leq \frac{x}{2} \\ r \equiv 1 \bmod p}} \frac{1}{r\log\frac{x}{r}} &= \frac{\pi(\frac{x}{2};p,1)}{\frac{x}{2}\log 2} + \int_{\frac{3p}{2}}^{\frac{x}{2}} \pi(t;p,1)\Big(\frac{1}{t^2\log\frac{x}{t}} - \frac{1}{t^2(\log\frac{x}{t})^2}\Big)\,dt \\
&\leq \frac{4}{p\log\frac{x}{2p}\log 2} + \int_{\frac{3p}{2}}^{\frac{x}{2}} \frac{4}{p\,t\log\frac{t}{p}\log\frac{x}{t}}\,dt \\
&= \frac{4}{p\log\frac{x}{2p}\log 2} + \frac{4}{p\log\frac{x}{p}}\int_{\frac{3p}{2}}^{\frac{x}{2}} \Big(\frac{1}{\log\frac{t}{p}} + \frac{1}{\log\frac{x}{t}}\Big)\frac{dt}{t} \\
&= \frac{4}{p\log\frac{x}{2p}\log 2} + \frac{4}{p\log\frac{x}{p}} \cdot \Big(\log\log\frac{t}{p} - \log\log\frac{x}{t}\Big)\Big|_{\frac{3p}{2}}^{\frac{x}{2}}.
\end{aligned}$$

Since $\frac{4}{\log 2 \log \frac{x}{2p}} < \frac{16}{\log \frac{x}{p}}$ and $16 - 4\log\log 2 - 4\log\log 3/2 < 22$, we see that

$$\sum_{\substack{r \le \frac{x}{2} \\ r \equiv 1 \bmod p}} \frac{1}{r \log \frac{x}{r}} \le \frac{1}{p \log \frac{x}{p}} \Big(22 + 8 \log\log \frac{2x}{3p}\Big)$$
$$\le \frac{30}{p \log \frac{x}{p}} \log\log x,$$

where for the last step we use $x \ge e^e$. □

This allows us to prove the following proposition:

Proposition 1. *For all $x \ge e^e$, each prime $p < x$, and all positive integers k, the number of primes $q \le x$ with $p \mid \varphi_k(q)$ is bounded above by*

$$\frac{4x}{p \log \frac{x}{p}} (31 \log\log x)^{k-1}.$$

Proof. We prove the proposition by induction. For $k = 1$, we have

$$\sum_{\substack{q \le x \\ p \mid \varphi(q)}} 1 = \pi(x; p, 1),$$

and the result follows from (2.1).

For $k \ge 1$, assume the proposition holds up to k. Then if $p \mid \varphi_{k+1}(q)$, either $p^2 \mid \varphi_k(q)$ or there exists a prime $r \equiv 1 \bmod p$ such that $r \mid \varphi_k(q)$. Thus,

$$\sum_{\substack{q \le x \\ p \mid \varphi_{k+1}(q)}} 1 \le \sum_{\substack{q \le x \\ p^2 \mid \varphi_k(q)}} 1 + \sum_{r \equiv 1 \bmod p} \sum_{\substack{q \le x \\ r \mid \varphi_k(q)}} 1.$$

In the second term, r is an odd prime and it divides the even number $\varphi_k(q)$, so $r \le \frac{1}{2}\varphi_k(q) \le \frac{1}{2}q \le \frac{1}{2}x$. Hence, by the induction hypothesis and Lemma 1, we have

$$\sum_{\substack{q \le x \\ p \mid \varphi_{k+1}(q)}} 1 \le \frac{4x}{p \log \frac{x}{p}} (31 \log\log x)^{k-1} + \sum_{\substack{r \le \frac{x}{2} \\ r \equiv 1 \bmod p}} \frac{4x}{r \log \frac{x}{r}} (31 \log\log x)^{k-1}$$
$$\le \frac{4x}{p \log \frac{x}{p}} (31 \log\log x)^{k-1} \Big(1 + 30 \log\log x\Big)$$
$$\le \frac{4x}{p \log \frac{x}{p}} (31 \log\log x)^{k},$$

where for the last inequality we use $x \ge e^e$. This proves the proposition. □

We now state a lemma and a proposition from [7]. Define $\Omega(n)$ to be the number of prime divisors of n, including multiplicity. We have the following lemma, the proof of which is suggested in Exercise 05 of [5]:

Lemma 2. *There exists an absolute constant $C_1 > 0$ such that for each positive integer k,*

$$\sum_{\substack{n \le x \\ \Omega(n) \ge k}} 1 \le C_1 \frac{k}{2^k} x \log x.$$

Define

$$F_K(n) = \prod_{0 \le k \le K} \varphi_k(n).$$

Proposition 2. *There exists an absolute constant $C_2 > 0$ such that for all sufficiently large numbers x, all numbers $y \ge 1$, and all integers $K \ge 1$, the number of integers $n \le x$ with $p^2 \mid F_K(n)$ for some prime $p > y$ is at most*

$$\frac{x}{y} K (C_2 \log\log x)^{2K}.$$

Clearly, the bound in the proposition still holds if we restrict the count to prime n. Using the arguments from [4] and [7] together with a short computation, we may take $C_2 = 27$.

We need one last proposition before moving on to the proof of the theorems.

Proposition 3. *There exists an absolute constant $C_3 > 0$ such that for each $x \ge e^e$ and every positive integer K, the number of primes $p \le x$ for which $\Omega(F_K(p)) > 2(6 \log\log x)^K$ is bounded above by*

$$C_3 \frac{x}{\log^2 x} (31 \log\log x)^{K-1}.$$

Proof. By Lemma 2, for any $t \in (1, x]$,

$$\begin{aligned} \sum_{\substack{q \le t \\ \Omega(q-1) > 6 \log\log x}} 1 &\le \sum_{\substack{n \le t-1 \\ \Omega(n) > 6 \log\log x}} 1 \le \frac{6\, C_1 \log\log x}{2^{6 \log\log x}}\, t \log t \\ &\le 6\, C_1 \frac{t \log\log x}{(\log x)^{6 \log 2 - 1}} \le 6\, C_1 \frac{t \,\log\log x}{(\log x)^3}. \end{aligned} \tag{2.2}$$

Since

$$6\, C_1 \frac{t \log\log x}{\log^3 x} \le 6\, C_1 \frac{x}{\log^2 x},$$

the proposition holds for $K = 1$, provided $C_3 \ge 6\, C_1$; we now assume $K \ge 2$.

If $\Omega(q-1) \le 6 \log\log x$ for each prime q dividing $F_{K-1}(p)$, then for all integers $1 \le k \le K$, $\Omega(\varphi_k(p)) \le (6 \log\log x)^k$ and $\Omega(F_K(p)) \le 2(6 \log\log x)^K$. Therefore, if $\Omega(F_K(p)) > 2(6 \log\log x)^K$, there exists a prime q dividing $F_{K-1}(p)$ with $\Omega(q-1) > 6 \log\log x$. Clearly, such a q must be odd, and since $F_{K-1}(p) = \prod_{k \le K-1} \varphi_k(p)$, $q \mid \varphi_k(p)$ for some $k \le K-1$. If this $k = 0$, then $q = p$; if $k > 0$, $\varphi_k(p)$ is even, so $q \le \frac{x}{2}$. Hence, we need only consider either $q = p$ or $q \le \frac{x}{2}$. Thus, by (2.2) and Proposition 1, we have that

$$\begin{aligned} \sum_{\substack{p \le x \\ \Omega(F_K(p)) > 2(6 \log\log x)^K}} 1 &\le \sum_{\substack{q \le x \\ \Omega(q-1) > 6 \log\log x}} 1 + \sum_{\substack{q \le \frac{x}{2} \\ \Omega(q-1) > 6 \log\log x}} \sum_{1 \le k \le K-1} \sum_{\substack{p \le x \\ q \mid \varphi_k(p)}} 1 \\ &\le \frac{6\, C_1\, x \,\log\log x}{\log^3 x} + \sum_{\substack{q \le \frac{x}{2} \\ \Omega(q-1) > 6 \log\log x}} \sum_{1 \le k \le K-1} \frac{4x (31 \log\log x)^{k-1}}{q \log \frac{x}{q}} \\ &\le \frac{6\, C_1\, x \,\log\log x}{\log^3 x} + \sum_{\substack{q \le \frac{x}{2} \\ \Omega(q-1) > 6 \log\log x}} \frac{5x (31 \log\log x)^{K-2}}{q \log \frac{x}{q}}. \end{aligned}$$

Approaching this last sum with partial summation, increasing the range of the sum to $q \le x$, and using the bound in (2.2), we have

$$\begin{aligned}\sum_{\substack{q \le \frac{x}{2} \\ \Omega(q-1) > 6 \log\log x}} \frac{1}{q \log \frac{x}{q}} &< \sum_{\substack{q \le x \\ \Omega(q-1) > 6 \log\log x}} \frac{2}{q} \\ &\le \frac{6\, C_1\, x\, \log\log x}{\log^3 x} \cdot \frac{2}{x} + \int_3^x \frac{6\, C_1\, t\, \log\log x}{\log^3 x} \cdot \frac{2}{t^2}\, dt \\ &\le \frac{12\, C_1\, \log\log x}{\log^2 x}.\end{aligned}$$

Putting this together, we have that

$$\begin{aligned}\sum_{\substack{p \le x \\ \Omega(F_K(p)) > 2(6 \log\log x)^K}} 1 &\le \frac{6\, C_1\, x \log\log x}{\log^3 x} + \frac{2\, C_1\, x\, (31 \log\log x)^{K-1}}{\log^2 x} \\ &\le \frac{6\, C_1\, x\, (31 \log\log x)^{K-1}}{\log^2 x},\end{aligned}$$

so we may take $C_3 = 6\, C_1$, which proves the proposition. □

3. The proof of Theorem 4

We are now ready to prove that $\varphi(F(p))$ grows faster than any fixed power of p for almost all primes p.

Proof. We follow the plan of Theorem 3 in [7]. Let $x > e^e$ and fix $0 < \epsilon < 1$ arbitrarily small. Set

$$K = \lceil (1-\epsilon) \log\log x / \log\log\log x \rceil.$$

For $p \le x$ prime, factor $F_K(p) = A \cdot B$, where A is the product of the primes dividing $F_K(p)$ which are less than or equal to $\log^3 x$, and B is the product of the primes larger than $\log^3 x$ (including multiplicity in both A and B). Proposition 3 gives that with at most $O(\frac{x}{\log^2 x}(31 \log\log x)^{K-1}) = o(\pi(x))$ exceptions, all choices of p have

$$A \le (\log^3 x)^{2(6 \log\log x)^K} = \exp((6 \log\log x)^{K+1}) = x^{o(1)}.$$

For x sufficiently large, the minimal order of $\varphi(m)/m$ for $m \le x$ (see [6, Theorem 328]) gives that $\varphi_{j+1}(p)/\varphi_j(p) > (2 \log\log x)^{-1}$ holds for each $j \in [0, K-1]$. Since $\pi(\frac{x}{2 \log\log x}) = o(\pi(x))$, we may assume $p > x/(2 \log\log x)$. Then,

$$\begin{aligned}F_K(p) = p^{K+1} \prod_{i=0}^{K} \frac{\varphi_i(p)}{p} &= p^{K+1} \prod_{i=1}^{K} \prod_{j=0}^{i-1} \frac{\varphi_{j+1}(p)}{\varphi_j(p)} \\ &> p^{K+1}/(2 \log\log x)^{1+2+\cdots+K} = p^{K+1}/(2 \log\log x)^{K(K+1)/2} \\ &> x^{K+1}/(2 \log\log x)^{(K+1)(K+2)/2} > x^{K+1/2}.\end{aligned}$$

So with $o(\pi(x))$ exceptions, all choices of $p \le x$ have

$$B = \frac{F_K(p)}{A} > x^{K+1/4}.$$

By Proposition 2, for sufficiently large x, the number of primes $p \le x$ such that $q^2 \mid F_K(p)$ for some $q \mid B$ is bounded above by

$$\begin{aligned}\frac{x}{\log^3 x}\frac{(\log\log x)(\log\log x)^{2K}}{\log\log\log x}\,(C_2)^{2K} &\le \frac{x}{\log^3 x}(\log\log x)^{2K+1}\,(C_2)^{2K}\\ &\le \frac{x}{\log^3 x}(\log x)^{2-2\epsilon}(\log\log x)^3\,(C_2)^{2K}\\ &= \frac{x}{(\log x)^{1+2\epsilon+o(1)}}\\ &= o(\pi(x)).\end{aligned}$$

So, but for $o(\pi(x))$ choices of the prime $p \le x$, B is squarefree. Since $B \mid F(p)$, we have that $\varphi(B) \mid \varphi(F(p))$. Again applying the minimal order of the Euler φ-function, we have

$$\varphi(B) > \frac{B}{2\log\log B} > \frac{x^{K+1/4}}{2(\log(K+1/4)+\log\log x)} > \frac{x^{K+1/4}}{3\log\log x} > x^K$$

for sufficiently large x. Thus, with at most $o(\pi(x))$ exceptions, $\varphi(F(p)) > x^K \ge p^K$ holds for primes $p \le x$. This completes the proof of Theorem 4. □

Corollary 2. *For almost all primes p, $F(p)$ grows faster than any fixed power of p.*

4. Lucas-Pratt trees

Now to attain our primary result, we will use Pratt's Theorem on $N(p)$, the number of primes appearing in the Lucas-Pratt tree for a prime p.

Proof. For each prime q appearing in the Lucas-Pratt tree for p, we have to at least show that $a^{\frac{q-1}{2}} \equiv -1 \bmod q$ for some primitive root a. This takes at least $\log_2(\frac{q-1}{2})$ modular multiplications $\bmod q$. Thus,

$$M(p) \ge \sum_{q\mid F(p)} \log_2\left(\frac{q-1}{2}\right) = \sum_{q\mid F(p)} \Big(\log_2(q-1)-1\Big) \ge \log_2\varphi(F(p)) - N(p).$$

Thus by Theorems 2 and 4 above, for each fixed $\epsilon > 0$, but for $o(\pi(x))$ primes $p \le x$,

$$\begin{aligned}M(p) &> \frac{1-\epsilon}{\log 2}\frac{\log p\log\log p}{\log\log\log p} - \frac{1}{\log 2}\log p\\ &= \frac{1}{\log 2}\Big((1-\epsilon)\frac{\log\log p}{\log\log\log p} - 1\Big)\log p.\end{aligned}$$

Since $\epsilon > 0$ is arbitrary, this completes the proof of Theorem 5. □

Acknowledgments

I would like to thank my advisor, Carl Pomerance, for his guidance on this work and extensive help in preparing this manuscript. I would also like to thank Hugh Williams for his helpful communication regarding Lucas's Theorem and an anonymous referee for some helpful comments.

References

1. M. Agrawal, N. Kayal, and N. Saxena, PRIMES is in P, *Annals of Mathematics*, 160 (2004), 781-793. MR2123939 (2006a:11170)
2. W. Banks and I. Shparlinski, On values taken by the largest prime factor of shifted primes, *Journal of the Australian Mathematical Society*, 82 (2007), 133–147.
3. R. Crandall and C. Pomerance, *Prime Numbers: A Computational Perspective*, Springer-Verlag, New York, 2001. MR1821158 (2002a:11007)
4. P. Erdős, A. Granville, C. Pomerance, and C. Spiro, On the normal behavior of the iterates of some arithmetic functions, in: *Analytic Number Theory, Proceedings of a Conference in Honor of P. T. Bateman*, Birkhäuser; Boston, Inc. 1990, 165-204. MR1084181 (92a:11113)
5. R. Hall and G. Tenenbaum, *Divisors*, Cambridge University Press, Cambridge, 1988. MR964687 (90a:11107)
6. G. Hardy and E. Wright, *An Introduction to the Theory of Numbers*, 4th ed., Clarendon Press, Oxford, UK, 1971. MR568909 (81i:10002)
7. F. Luca and C. Pomerance, Irreducible radical extensions and Euler-function chains, *Integers*, to appear, (Available at http://math.dartmouth.edu/∼carlp/PDF/radicalpaper5.pdf).
8. E. Lucas, Considérations nouvelles sur la théorie des nombres premiers et sur la division géométrique de la circonférence en parties égales, Association Française pour l'Avancement des Sciences, *Comptes Rendus*, 6 (1877), 159-167.
9. H. Montgomery and R. Vaughan, The large sieve, *Mathematika* (20) **40** (1973), 119-134. MR0374060 (51:10260)
10. C. Pomerance, Very short primality proofs, *Mathematics of Computation* (48) **177** (1987), 315-322. MR866117 (88b:11088)
11. V. Pratt, Every prime has a succinct certificate, *SIAM Journal on Computing*, 4 (1975), 214-220. MR0391574 (52:12395)

Department of Mathematics, Dartmouth College, Hanover, New Hamphshire 03755-3551

E-mail address: jonathan.bayless@dartmouth.edu