MATHEMATICS OF COMPUTATION
Volume 74, Number 252, Pages 2035–2041
S 0025-5718(05)01749-7
Article electronically published on February 16, 2005

# ON THE GREATEST PRIME FACTOR OF $p-1$ WITH EFFECTIVE CONSTANTS

G. HARMAN

ABSTRACT. Let $p$ denote a prime. In this article we provide the first published lower bounds for the greatest prime factor of $p-1$ exceeding $(p-1)^{\frac{1}{2}}$ in which the constants are effectively computable. As a result we prove that it is possible to calculate a value $x_0$ such that for every $x > x_0$ there is a $p < x$ with the greatest prime factor of $p-1$ exceeding $x^{\frac{3}{5}}$. The novelty of our approach is the avoidance of any appeal to Siegel's Theorem on primes in arithmetic progression.

## 1. INTRODUCTION

The problem of finding good lower bounds for the greatest prime factor of $p-1$ where $p$ is a prime or, more generally, of $p+a$ for some fixed integer $a$ has attracted great attention from a variety of areas. The cases $a = -1$ and 2 are the best known, being approximations for the widely believed conjectures that for infinitely many primes $q$ we have $2q+1$ also prime and the "twin-prime conjecture": for infinitely many primes $p$ the expression $p+2$ is also prime. Although number-theorists have always had an interest in such problems, in recent years cryptographers, group-theorists, and computer scientists have had occasion to apply such results in their own fields. Most recently the work of Agrawal, Kayal and Saxena [1] has given much publicity to the case $p-1$. Although subsequent proofs have been given which do not depend on results from analytic number theory [4], the original version of the proof that "PRIMES is in P" required a result of the form

$$P(p-1) > p^{\theta},$$

with $\theta > \frac{1}{2}$, where we use $P(n)$ to denote the greatest prime factor of an integer $n$. The first result of this type was due to Goldfeld [9] who obtained $\theta = 0.583\cdots$. Since that time there have been many improvements in the value for $\theta$ culminating in the work of Baker and Harman [2, 3] who proved

**Theorem 1.** *Let $a \in \mathbb{N}$. For $\theta \le 0.677$ and all large $x > X_0(a,\theta)$ we have*

$$\sum_{\substack{p\le x\\ P(p+a)>x^{\theta}}} 1 > \delta(\theta)\frac{x}{\log x}, \tag{1.1}$$

*where $\delta(\theta) > 0$.*

Received by the editor March 19, 2004 and, in revised form, August 16, 2004.
2000 *Mathematics Subject Classification.* Primary 11N13.

Although not stated with an explicit lower bound, it is clear that the method gives such a term on the right of (1.1) (compare our argument below leading to (2.4)). This result, like all its antecedents, is ineffective in the sense that we have no way of calculating the constant $X_0(a,\theta)$. The purpose of this note is to prove an effective form of a weaker result. We will explain why all the constants involved in the proof are effective, but we will make no attempt to calculate them. There are various tricks that can be applied at various points by splitting up integration ranges and using different inequalities in each range, and using the fact that we know there are no Siegel zeros to small moduli, and so reduce the values of the constants obtained. The purpose of our work is to show that the constants could be calculated (that is, the calculation is *feasible* as well as effective), against a common impression that analytic number theory can only give results with unspecified constants. A reader with a few days to spare and a computational aptitude should have no trouble in working out a permissible value of the constant for $\theta = 0.6, a = -1$, say. Our main theorem is as follows.

**Theorem 2.** *Let $a \in \mathbb{Z}, \theta < 1 - \frac{1}{2}\exp\left(-\frac{1}{4}\right) = 0.6105\cdots$. Then there exist effectively computable constants $X_1(a,\theta), \delta(\theta) > 0$, such that, if $x > X_1$, we have*

$$\sum_{\substack{p \le x \\ P(p+a) > x^\theta}} 1 \; > \; \delta(\theta)\frac{x}{\log x}. \tag{1.2}$$

We note that a theorem with this exponent was first proved by Motohashi [11], although he used a result with ineffective constants. It may well be possible to improve the value of $\theta$ while keeping all the constants effective, but the results of Baker and Harman appear inaccessible in this situation. Indeed, using the results of Deshouillers and Iwaniec [7], one should obtain $\theta < 1 - \frac{1}{2}\exp\left(-\frac{3}{8}\right)$, so that $\theta = 0.65$ becomes permissible. However, the justification of the effectiveness of the constants and the practicality of calculating them in that case appear far from simple. We remark that the dependence of $X_1$ on $a$ is quite mild—the result could be made uniform in $|a| < x^b$ where $b < 1$ is fixed.

## 2. The method discussed

We write $\pi(x,b,q)$ for the number of primes up to $x$ which are congruent to $b \bmod q$, and we use $\Lambda(n), \phi(n)$ for von Mangoldt's function and Euler's totient function, respectively. We shall take it for granted that the error term in the Prime Number Theorem, and its repercussions for sums over primes converted by partial summation to integrals, can be calculated with explicit constants; see [8] for example.

As is well known to workers in number theory, ineffective constants are usually hard to avoid when discussing primes in arithmetic progressions in view of Siegel's Theorem (see [6, Chapter 21]). Previous workers have appealed to the Bombieri-Vinogradov theorem which implicitly uses that result. The Baker-Harman proof also appeals to deep results of Bombieri, Friedlander and Iwaniec (see [5] for example), and it is not clear how the constants arising from this method can be made effective. We shall demonstrate that an effective variant of the Bombieri-Vinogradov Theorem can be employed and this leads to our result. Before outlining the proof, we state an important lemma (see [10] for the proof).

**Lemma 1** (The Brun-Titchmarsh Inequality)**.** *For all $1 \le q < x$ and all $b \in \mathbb{N}$ we have*

$$\pi(x,b,q) < \frac{2x}{\phi(q)\log(x/q)}. \tag{2.1}$$

*Remark.* It is generally conjectured that

$$\pi(x,b,q) \sim \frac{x}{\phi(q)\log(x)}$$

over a large range of $q$, but all we are able to do is show that this holds on average over a certain range (see Theorem 3 below). Better bounds than (2.1) can be proved for the parameters in certain ranges using deep results on averages of Kloosterman sums, and this leads to problems when trying to calculate the constants involved.

We begin the proof as previous authors have done with a device invented by Chebychev. We assume throughout that $a = -1$ for convenience. We have

$$\begin{aligned}\sum_{p_1\le x}\sum_{\substack{p_2\mid(p_1-1)\\ p_2>N}}\log p_2 &= x - \sum_{p_2<N}(\log p_2)\pi(x,1,p_2) \;+\; o(x)\\ &= x - U(x,N) + o(x),\end{aligned} \tag{2.2}$$

say. We have

$$\sum_{\substack{p_2\mid(p_1-1)\\ p_2>N}}\log p_2 \le \begin{cases}\log x & \text{if } P(p_1-1)>N,\\ 0 & \text{otherwise.}\end{cases}$$

Hence, if we show that

$$U(x,N) < (1-\eta)x \tag{2.3}$$

for some $\eta > 0$, then

$$\sum_{\substack{p\le x\\ P(p-1)>N}} 1 \;>\; \frac{(\eta-o(1))x}{\log x} \;>\; \frac{\eta x}{2\log x}, \tag{2.4}$$

for all sufficiently large $x$. In (2.2) we have our first implied constant in the $o(x)$ term, so we must justify that this formula contains effective constants. In detail the derivation of the formula goes like this:

$$\begin{aligned}\sum_{p_1\le x}\sum_{p_2\mid(p_1-1)}\log p_2 &= \sum_{p_1\le x}\sum_{n\mid(p_1-1)}\Lambda(n) - \sum_{p_1\le x}\sum_{\substack{p^k\mid(p_1-1)\\ k\ge 2}}\log p\\ &= \sum_{p_1\le x}\log(p_1-1) - E_1,\end{aligned}$$

say, using the well-known formula $\sum_{d\mid n}\Lambda(d) = \log n$

$$\begin{aligned}&= \sum_{n\le x}\Lambda(n) - \sum_{\substack{p^k\le x\\ k\ge 2}}\log p + \sum_{p_1\le x}\log\left(1-\frac{1}{p_1}\right) - E_1\\ &= x + E_4 - E_3 - E_2 - E_1,\end{aligned}$$

say, where we have written $E_4$ for the error term from the Prime Number Theorem:

$$\sum_{n\le x}\Lambda(n) = x + E_4.$$

Now there are no problems giving a suitable bound on $E_4$ [8]. The term $E_3$ does not exceed $2x^{\frac12}$ for all large $x$. The term $E_2$ is practically trivial, being no more than $\log x$ for $x \ge 2$. The problematic term is $E_1$. We have

$$\begin{aligned} E_1 &= \sum_{k\ge2}\sum_{p\le x^{\frac1k}} (\log p)\pi(x,1,p^k) \\ &\le \sum_{k\ge2}\sum_{p\le x^{\frac1k}} (\log p)\min\left(\frac{2x}{\phi(p^k)\log(x/p^k)}, \frac{x}{p^k}\right) \\ &\le \frac{4x}{\log x}\sum_{k\ge2}\sum_{p\le x^{\frac{1}{2k}}}\frac{\log p}{\phi(p^k)} + \sum_{k\ge2}\sum_{x^{\frac{1}{2k}}<p\le x^{\frac1k}}(\log p)\frac{x}{p^k} \\ &\le \frac{4x}{\log x}\sum_{p=2}^{\infty}\frac{\log p}{(p-1)^2} + x^{\frac34}\log x + x^{\frac23}(\log x)^2, \end{aligned}$$

with some simple upper bounds. We note that

$$\sum_{p=2}^{\infty}\frac{\log p}{(p-1)^2} < \frac54,$$

so that $E_1 < 5\frac{x}{\log x}$ plus much smaller terms. To obtain an exponent of $\frac35$, we would need to take $x > \exp(500)$ at least to get this term suitably small. Although more careful working can reduce the value 5, this term may well represent one of the main errors. There is a stronger competitor for this dubious honour later on in the argument, however.

Our task is now to bound $U(x,N)$. If we applied (2.1) for every single $p_2 < N = x^v$, we would obtain (not bothering to write in the errors explicitly)

$$U(x,N) < 2x\int_0^v \frac{1}{1-\alpha}\,d\alpha + \cdots = 2x\log\left(\frac{1}{1-v}\right) + \cdots < (1-\eta)x$$

if $v < 1-\exp(-\frac12) = 0.393\cdots$ and $x$ is suffciently large. Our plan is to apply (2.1) only when the Bombieri-Vinogradov Theorem fails (this was Motohashi's argument) or for small $p_2$ where a potential Siegel zero could disrupt the distribution of primes.

## 3. The Bombieri-Vinogradov Theorem revisited

The Bombieri-Vinogradov Theorem states that $\pi(x,b,q)$ is $\mathrm{Li}(x)/\phi(q)$ on average over $q$ almost up to $x^{\frac12}$, where

$$\mathrm{Li}(x) = \int_2^x \frac{1}{\log y}\,dy.$$

To be precise, we have the following ([6, Chapter 28] after partial summation).

**Theorem.** *Let $x \ge 2, A > 0$. Write $X = \sqrt{x}, \mathcal{L} = \log x$. Then we have*

$$\sum_{q<X\mathcal{L}^{-A}}\max_{(a,q)=1}\left|\pi(x,a,q) - \frac{\mathrm{Li}(x)}{\phi(q)}\right| \ll_A x\mathcal{L}^{5-A}.$$

*Remark.* To be nontrivial, one needs to take $A > 5$ in the above, of course. The problem with the result is the implied constant in $\ll_A$ which is ineffective.

We now give a variant of the above result which has no ineffective constants, but only allows prime moduli. We remark that independently Lenstra and Pomerance (work in preparation on primality testing with Gaussian periods) have given a similar result. Timofeev [12] has also given an effective version of Theorem 2, but his result involves a reference to a possible exceptional zero. To be precise, the right-hand side of the inequality in his theorem involves a term

$$\frac{x^\beta}{\phi(k)}\mathcal{L}^{\frac{5}{4}},$$

where $k$ is a modulus having exceptional zero $\beta$. Our approach removes any such reference.

**Theorem 3.** *Let $x \geq 2$. Suppose that $1 \leq U < V \leq x^{\frac{1}{2}}$. Write $\mathcal{L} = \log x$. Then, with $p$ denoting a prime variable,*

$$\sum_{U \leq p \leq V} \max_{(a,p)=1} \left|\pi(x,a,p) - \frac{\pi(x)}{\phi(p)}\right| \ll \left(\frac{x}{U} + x^{\frac{5}{6}} + x^{\frac{1}{2}}V\right)\mathcal{L}^5 \tag{3.1}$$

*where the implied constant is absolute and effectively computable.*

*Remark.* If one replaces $\pi(x)$ by $\mathrm{Li}(x)$ in (3.1), then a term like $x\exp\left(-(\log x)^{\frac{1}{2}}\right)$ must be added on the right-hand side, but the constants remain effectively computable since this is just the error in the prime number theorem times

$$\sum_{U \leq p \leq V} \frac{1}{p-1}.$$

*Proof.* This is essentially established in [6, Chapter 28] (Vaughan's elementary proof of the Bombieri-Vinogradov Theorem originally given in [13]). We need only note that to a prime modulus all nonprincipal characters are primitive (and this relieves us of the messy task of allowing for a potential small bad modulus which can divide larger moduli) and so the bound in the display after (3) on page 164 of [6] holds for all such characters. That the implied constants are absolute and effectively computable may be discovered by careful scrutiny of the working. We need only determine the implied constants in simple inequalities such as (next to the last display on page 166 of [6])

$$\sin(t\log u) \ll \min(1, |t|\log(2MN))$$

if $u = k + \frac{1}{2}$ where $k \in \mathbb{Z} \cap [0, MN]$ (the constant is in fact 1 here), or in standard number theory bounds such as

$$\sum_{n \leq N} \tau^2(n) \ll N\log^3 N, \tag{3.2}$$

for $N \geq 2$. In fact, for (3.2) we have a clean inequality with constant 1 true for $N \geq 1$ if we replace the right-hand side by $N(\log(eN))^3$. Indeed, all the implied constants in the proof are relatively small. □

## 4. Completion of the proof

Let $1 < U < V < x^{\frac{1}{2}}$. We apply (2.1) for $p_2 < U$ and for $p_2 > V$. The region $U \le p_2 \le V$ will be covered by (3.1). Suppose we take $U = x^{\frac{1}{2}}/V = \mathcal{L}^{10}$, say. We make no attempt here to optimise the choice for these parameters. Then

$$\sum_{2\le p<U} (\log p)\pi(x,1,p) < \sum_{2\le p<U} \frac{2x\log p}{(p-1)\log(x/p)} < 30x\frac{(\log\log x)}{\log x},$$

for all large $x$. There would be no problem calculating a suitable lower bound for $x$ to make this inequality work. Also, if $N = x^\alpha$, we have

$$\begin{aligned}\sum_{V<p\le N} (\log p)\pi(x,1,p) &< \sum_{V<p\le N} \frac{2x\log p}{p\log(x/p)} + O\left(\frac{x}{V}\right)\\ &= 2x\int_{\frac{1}{2}}^{\alpha} \frac{1}{1-u}\,du + O\left(x\frac{\log\log x}{\log x}\right)\\ &= 2x\log\left(\frac{1}{2(1-\alpha)}\right) + O\left(x\frac{\log\log x}{\log x}\right).\end{aligned}$$

Finally, applying (3.1) for $U \le p \le V$, we obtain

$$\sum_{U\le p\le V} (\log p)\pi(x,1,p) < \frac{x}{2},$$

for all large $x$, again with an effectively computable lower bound for $x$ in view of the effectively computable constants in (3.1). Altogether we thus arrive at the bound

$$U(x,N) < x\left(\frac{1}{2} + 2\log\left(\frac{1}{2(1-\alpha)}\right)\right) + O\left(\frac{x\log\log x}{\log x}\right),$$

which establishes (2.3) if $\alpha < 1 - \frac{1}{2}\exp\left(-\frac{1}{4}\right)$ and so completes the proof.

It should be noted that the terms which give rise to errors of the form $x\frac{\log\log x}{\log x}$ (which appear to be the largest errors encountered) come not from any deep analyic number theory, but simply from summing over all small $p$ separately and from replacing

$$\frac{1}{2} - 10\frac{\log\log x}{\log x} \quad \text{by} \quad \frac{1}{2}.$$

ACKNOWLEDGMENT

The author thanks the referee for his comments and Professor Carl Pomerance for interesting e-mail correspondence on this topic.

## References

1. M. Agrawal, N. Kayal and N. Saxena, *PRIMES is in P*, http://www.cse.iitk.ac.in/primality.pdf.
2. R. C. Baker, G. Harman, *The Brun-Titchmarsh theorem on average*, Analytic Number Theory (Proceedings in honor of Heini Halberstam), Birkhauser, Boston, 1996, 39-103. MR1399332 (97h:11096)
3. R. C. Baker, G. Harman, *Shifted primes without large prime factors*, Acta Arith. **83** (1998), 331-361. MR1610553 (99b:11104)
4. D. Bernstein, *Proving primality after Agrawal-Kayal-Saxena*, http://cr.yp.to/papers/html#aks.
5. E. Bombieri, J. B. Friedlander and H. Iwaniec, *Primes in arithmetic progressions to large moduli III*, J. American Math. Soc. **2** (1989), 215-224. MR0976723 (89m:11087)

6. H. Davenport, *Multiplicative Number Theory* (second edition revised by H. L. Montgomery), Springer-Verlag, New York, 1980. MR0606931 (82m:10001)
7. J.-M. Deshouillers and H. Iwaniec, *On the Brun-Titchmarsh theorem on average* in Topics in classic number theory (ed. G. Halász), vol. 1 (Budapest, 1981), 319-333. MR0781145 (86e:11085)
8. K. Ford, *Vinogradov's Integral and bounds for the Riemann zeta-function*, Proc. London Math. Soc. (3) **85** (2002), 565-633. MR1936814 (2003j:11089)
9. M. Goldfeld, *On the number of primes p for which $p+a$ has a large prime factor*, Mathematika **16** (1969), 23-27. MR0244176 (39:5493)
10. H. L. Montgomery and R. C. Vaughan, *The large sieve*, Mathematika **20** (1973), 119-134. MR0374060 (51:10260)
11. Y. Motohashi, *A note on the least prime in an arithmetic progression with a prime difference*, Acta Arith. **17** (1970), 283-285. MR0268131 (42:3030)
12. N. M. Timofeev, *The Vinogradov-Bombieri theorem*, (English) Math. Notes **38** (1985), 947-951. MR0823418 (87f:11073)
13. R. C. Vaughan, *An elementary method in prime number theory*, Acta Arith. **37** (1980), 111-115. MR0598869 (82c:10055)

Department of Mathematics, Royal Holloway University of London, Egham, Surrey TW20 0EX, United Kingdom

*E-mail address*: G.Harman@rhul.ac.uk