

ANNALES DE L'INSTITUT FOURIER

JEAN-MARC DESHOUILLERS

HENRYK IWANIEC

On the greatest prime factor of $n^2 + 1$

Annales de l'institut Fourier, tome 32, n° 4 (1982), p. 1-11

http://www.numdam.org/item?id=AIF_1982__32_4_1_0

© Annales de l'institut Fourier, 1982, tous droits réservés.

L'accès aux archives de la revue « Annales de l'institut Fourier » (<http://annalif.ujf-grenoble.fr/>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques

<http://www.numdam.org/>

ON THE GREATEST PRIME FACTOR OF $n^2 + 1$

by J.-M. DESHOILLERS and H. IWANIEC

1. Introduction.

In 1967 C. Hooley [2] (see also [3]) showed that if D is not a perfect square then the greatest prime factor of $n^2 - D$ exceeds $n^{11/10}$ infinitely often. In fact Hooley's arguments yield a slightly better result with the exponent $11/10$ replaced by any θ less than $\theta_0 = 1.100\ 148\ 3 \dots$ the solution of

$$(1) \quad \frac{14}{3} \left(\theta - \frac{12}{11} \right) + \frac{28}{9} \log \left(1 + \frac{33}{14} \left(\theta - \frac{12}{11} \right) \right) - \frac{41}{33} + \frac{32}{9} \log \frac{11}{8} = 0.$$

Among several innovative ideas in Hooley's proof one finds a very interesting application of A. Weil's estimate for Kloosterman sums

$$(2) \quad S(nm; c) = \sum_{\substack{d \pmod{c} \\ (d, c) = 1}} e \left(n \frac{\bar{d}}{c} + m \frac{d}{c} \right) \ll (n, m, c)^{\frac{1}{2}} c^{\frac{1}{2} + \varepsilon}$$

where the symbol \bar{d} stands for a solution of $d\bar{d} \equiv 1 \pmod{c}$. Recently the authors [1] investigated linear forms in Kloosterman sums $S(nQ, m; c)$ with the variables of the summation n, m and c counted with a smooth weight function, showing (see Lemma 3) that there exists a considerable cancellation of terms.

In the paper we inject this result into the Chebyshev-Hooley method to prove the following

THEOREM. — *For any $\varepsilon > 0$ there exist infinitely many integers n such that $n^2 + 1$ has a prime factor greater than $n^{\theta - \varepsilon}$, where θ satisfies*

$$2 - \theta - 2 \operatorname{Log} (2 - \theta) = \frac{5}{4} \quad (\theta = 1.202\ 468 \dots).$$

Our result can be generalized to $n^2 - D$ by using Hooley's arguments from [3].

The authors express their thanks to Prof. C. Hooley for interesting comments and corrections.

2. Chebyshev's method.

Let $x \geq 2$ and let b be a non-negative function of C^∞ -class with support in $[x, 2x]$ and the derivatives of which satisfy

$$b^{(l)}(\xi) \ll x^{-l}, \quad l = 0, 1, 2, \dots,$$

the implied constant in \ll depending on l alone. Denote

$$X = \int b(\xi) d\xi \quad \text{and} \quad |A_d| = \sum_{n^2+1 \equiv 0 \pmod{d}} b(n).$$

We begin with applying Chebyshev's idea to calculate

$$\begin{aligned} (1) \quad T(x) &= \sum_p |A_p| \log p = \sum_d |A_d| \wedge(d) + O(x) \\ &= \sum_n b(n) \sum_{d|n^2+1} \wedge(d) + O(x) = \sum_n b(n) \log(n^2+1) + O(x) \\ &= 2(\log x) \int b(\xi) d\xi + O(x) = 2X \log x + O(x). \end{aligned}$$

The partial sum

$$\begin{aligned} T_0(x) &= \sum_{p \leq x} |A_p| \log p \\ &= \sum_{p \leq x} \sum_{v^2+1 \equiv 0 \pmod{p}} (\log p) \sum_{n \equiv v \pmod{p}} b(n) \end{aligned}$$

can be evaluated easily by the Poisson summation formula.

LEMMA 1. — For any $f(\xi)$ of C^1 class with compact support in $(0, \infty)$ we have

$$\sum_{n \equiv a \pmod{q}} f(n) = \frac{1}{q} \sum_{h \pmod{q}} e\left(-\frac{ah}{q}\right) \hat{f}\left(\frac{h}{q}\right), \quad h \in \mathbf{Z}$$

where $\hat{f}(t)$ is the Fourier transform of $f(\xi)$.

By Lemma 1

$$\sum_{n \equiv v \pmod{p}} b(n) = \frac{1}{p} \sum_h e\left(-\frac{vh}{p}\right) \hat{b}\left(\frac{h}{p}\right).$$

For $h = 0$ we have $\hat{b}(0) = X$. If $h \neq 0$ by partial integration two times we get

$$\begin{aligned} \hat{b}\left(\frac{h}{p}\right) &= \int b(\xi) e\left(\frac{h}{p} \xi\right) d\xi \\ &= \left(\frac{p}{2\pi i h}\right)^2 \int b''(\xi) e\left(\frac{h}{p} \xi\right) d\xi \ll h^{-2} p^2 x^{-1}. \end{aligned}$$

This yields

$$\sum_{n \equiv v \pmod{p}} b(n) = \frac{X}{p} + O\left(\frac{p}{x}\right)$$

whence

$$(2) \quad T_0(x) = X \log x + O(x).$$

Letting P_x be the greatest prime factor of $\prod_{x < n < 2x} (n^2 + 1)$ by (1) and (2) it follows that

$$(3) \quad S(x) = \sum_{x < p \leq P_x} |\mathcal{A}_p| \log p = X \log x + O(x).$$

Our aim is to estimate $S(x)$ from above and deduce from it a lower estimate for P_x .

3. Splitting up of $S(x)$.

In what follows it will be convenient to have p counted with a smooth weight function. Therefore we arrange the sum $S(x)$ as

$$(4) \quad S(x) = \sum_{1 \leq j \leq J} S(x, P_j) + O(x)$$

with $P_j = 2^j x$, $0 \leq j \leq J \leq 2 \log x$ and

$$(5) \quad S(x, P_j) = \sum_{P_j < p \leq 4P_j} |\mathcal{A}_p| C_j(p) \log p$$

where $C_j(\xi)$ are non-negative functions of C^∞ -class which satisfy the three following conditions

$$\text{Supp } C_j \subset [P_j, 4P_j]$$

$$(6) \quad \sum_{0 \leq j \leq J} C_j(\xi) = \begin{cases} 1 & \text{if } 2x < \xi \leq P_x \\ O(1) & \text{if } x < \xi \leq 2x \text{ or } P_x < \xi \leq 2P_x \\ 0 & \text{otherwise} \end{cases}$$

$C_j^{(l)}(\xi) \ll P_j^{-l}$, with the implied constant depending on l alone. The error term $O(x)$ in (4) comes from a trivial estimate for the contribution of primes p in the interval $(x, 2x]$ which is not completely covered.

4. Application of the sieve method.

A typical sum to be considered is

$$S(x, P) = \sum_{P < p \leq 4P} |\mathcal{A}_p| C(p) \log p$$

with $x < P \leq 2P_x$. Let $x \geq D \geq 1$ and let $\{\lambda_d\}_{d \leq D}$ be an upper bound sieve of level D , i.e. a sequence of real numbers such that

$$\lambda * 1 \geq \mu * 1, \quad \lambda_1 = 1, \quad \lambda_d = 0 \text{ for } d \geq D.$$

We also assume that $|\lambda_d| \leq 1$ for all d and that $\lambda_d = 0$ when d is not square-free.

Thus

$$S(x, P) \leq \sum_{d \leq D} \lambda_d \sum_{m \equiv 0 \pmod{d}} |\mathcal{A}_m| C(m) \log m.$$

By the Poisson summation formula we write

$$\begin{aligned} |\mathcal{A}_m| &= \sum_{v^2 + 1 \equiv 0 \pmod{m}} \sum_{n \equiv v \pmod{m}} b(n) \\ &= \frac{\omega(m)}{m} X + r(\mathcal{A}, m) \end{aligned}$$

where $\omega(m)$ is the number of incongruent solutions of $v^2 + 1 \equiv 0 \pmod{m}$ and

$$(7) \quad r(\mathcal{A}, m) = \frac{1}{m} \sum_{h \neq 0} \sum_{v^2 + 1 \equiv 0 \pmod{m}} e\left(-\frac{vh}{m}\right) \hat{b}\left(\frac{h}{m}\right).$$

According to the above we write

$$(8) \quad S(x, P) \leq XV(x, P) + R(x, P)$$

where $XV(x, P)$ is considered as a main term

$$V(x, P) = \sum_{d < D} \lambda_d \sum_{m \equiv 0 \pmod{d}} \frac{\omega(m)}{m} C(m) \log m$$

and $R(x, P)$ is the total error term

$$R(x, P) = \sum_{d < D} \lambda_d R(x, d, P)$$

with

$$(9) \quad R(x, d, P) = \sum_{h \neq 0} \sum_{m \equiv 0 \pmod{d}} \frac{C(m)}{m} \log m \sum_{v^2 + 1 \equiv 0 \pmod{m}} \hat{b}\left(\frac{h}{m}\right) e\left(-\frac{vh}{m}\right).$$

5. Transformation of $R(x, d, P)$.

We are searching for D as large as possible for which the estimate

$$(10) \quad R(x, P) \ll x^{1-\epsilon}$$

is available. By partial integration $l = [4\epsilon^{-1}]$ times we get

$$\hat{b}\left(\frac{h}{m}\right) = \left(-2\pi i \frac{h}{m}\right)^{-l} \int b^{(l)}(\xi) e\left(\frac{h}{m} \xi\right) d\xi \ll x \left(\frac{P}{|h|x}\right)^l \ll h^{-2}$$

for $|h| \geq Px^{\epsilon-1} = H$, say. Hence truncating the series (9) at $h = H$ we make an error $O(\tau(d)/d)$ which contributes to $R(x, P)$ an admissible amount

$$\sum_{d < D} \frac{\tau(d)}{d} \ll (\log D)^2 \ll (\log x)^2.$$

For the remaining terms we need an explicit formula for the solutions of

$$(11) \quad v^2 + 1 \equiv 0 \pmod{m}.$$

LEMMA 2. (Gauss). — Let $m > 1$. If (11) is soluble then m is represented properly as a sum of two squares

$$(12) \quad m = r^2 + s^2, \quad (r, s) = 1, \quad r, s > 0.$$

There is a one to one correspondence between the incongruent solutions $v \pmod{m}$ of (11) and the solutions (r, s) of (12) given by

$$\frac{v}{m} = \frac{\bar{r}}{s} - \frac{r}{s(r^2 + s^2)}.$$

Proof. — See [5] and [3], p. 34, eq. (68).

By Lemma 2 we get

$$\sum_{v^2+1 \equiv 0 \pmod{m}} e\left(-\frac{vh}{m}\right) = \sum_{\substack{r^2+s^2=m \\ r,s>0, (r,s)=1}} e\left(-h\frac{\bar{r}}{s}\right) \left\{1 + O\left(\frac{r|h|}{sm}\right)\right\}$$

whence letting $g(m, h) = \frac{C(m)}{m} (\log m) \hat{b}\left(\frac{h}{m}\right)$ we obtain

$$R(x, d, P) = \sum_{0 < |h| \leq H} \sum_{\substack{(r,s)=1, r,s>0 \\ r^2+s^2 \equiv 0 \pmod{d}}} g(r^2+s^2, h) e\left(-h\frac{\bar{r}}{s}\right) + O(d^{-1} P x^{3\epsilon-1}).$$

Here the error $O(d^{-1} P x^{3\epsilon-1})$ contributes to $R(x; P)$ less than $P x^{3\epsilon-1} \log x \ll x^{1-\epsilon}$ provided $P \leq x^{2-5\epsilon}$ which we henceforth assume.

For sum over r we apply Poisson's summation formula giving

$$\begin{aligned} & \sum_{\substack{(r,s)=1 \\ r^2+s^2 \equiv 0 \pmod{d}}} g(r^2+s^2, h) e\left(-h\frac{\bar{r}}{s}\right) \\ & \quad \sum_{\substack{u \pmod{ds} \\ (u,s)=1 \\ u^2+s^2 \equiv 0 \pmod{d}}} e\left(-h\frac{\bar{u}}{s}\right) \sum_{r \equiv u \pmod{ds}} g(r^2+s^2, h) \\ & = \frac{1}{ds} \sum_k \sum_{\substack{u \pmod{ds} \\ (u,s)=1 \\ u^2+s^2 \equiv 0 \pmod{d}}} e\left(-h\frac{\bar{u}}{s} - k\frac{u}{ds}\right) G(h, k; s) \end{aligned}$$

where $G(h, k; s) = \int g(\xi^2 + s^2, h) e(k\xi/ds) d\xi$. Writing $u = \alpha s + \beta d$ with $\alpha^2 + 1 \equiv 0 \pmod{d}$ it becomes

$$\frac{1}{ds} \sum_{\alpha^2+1 \equiv 0 \pmod{d}} \sum_k e\left(-\frac{\alpha k}{d}\right) S(-h\bar{d}, -k; s) G(h, k; s).$$

For $k = 0$ the Kloosterman sum $S(-h\bar{d}, -k; s)$ reduces to a Ramanujan sum for which we have

$$|S(-h\bar{d}, 0; s)| \leq (h, s).$$

Therefore the terms with $k = 0$ contribute less than

$$\frac{\tau(d)}{d} \sum_{0 < |h| \leq H} \sum_{s \leq 2\sqrt{P}} \frac{(h, s) x \text{Log } P}{s P} \ll \frac{\sqrt{P}}{d} x^\epsilon \ll \frac{x^{1-\epsilon}}{d}.$$

Finally

$$(13) \quad R(x, d, P) = \frac{1}{d} \sum_{\alpha^2 + 1 \equiv 0 \pmod{d}} \sum_{0 < |h| \leq H} \sum_{k \neq 0} e\left(-\frac{\alpha k}{d}\right) \sum_{\substack{s > 0 \\ (s, d) = 1}} \frac{1}{s} S(-h\bar{d}, -k; s) G(h, k; s) + O\left(\frac{x^{1-\epsilon}}{d}\right).$$

6. Linear forms in Kloosterman sums.

Let $N, M, C \geq 1$ and $f(n, m, c)$ be a function of C^6 class with compact support in $[C, 2C]$ with respect to c and satisfying

$$(14) \quad \left| \frac{\partial^{l_1 + l_2 + l_3}}{\partial n^{l_1} \partial m^{l_2} \partial c^{l_3}} f(n, m, c) \right| \leq N^{-l_1} M^{-l_2} C^{-l_3}, \quad 0 \leq l_1, l_2, l_3 \leq 2.$$

In this section we borrow from [1] an estimate for the average of trilinear forms (cf. *Theorem 11*)

$$B_d^\pm(N, M, C) = \sum_{0 < n \leq N} \sum_{0 < m \leq M} \sum_{(c, d) = 1} b_{m, d} S(n\bar{d}, \pm m; c) f(n, m, c)$$

where $b_{m, d}$ are arbitrary complex numbers.

LEMMA 3. — If $f(n, m, c)$ satisfies (14) then for any $\epsilon > 0$ we have

$$\left(\sum_{D < d \leq 2D} |B_d^\pm(N, M, C)| \right)^2 \ll (CDMN)^\epsilon N \left(\sum_{\substack{0 < m \leq M \\ D < d \leq 2D}} |b_{m, d}|^2 \right) \times \left\{ \frac{D(DC^2 + MN + NC^2)(DC^2 + MN + MC^2)}{DC^2 + MN} + \leq \sqrt{D(D+M)} \cdot C^3 \right\}$$

and the same upper bound holds for $(\sum |B_d^-|)^2$, the constant implied in \ll depending on ϵ at most.

7. Estimation of the error.

In order to make Lemma 3 applicable we first split up the sum over s in $R(x,d,P)$ into $\ll \log P$ sums of the type

$$(14) \quad \sum_{(s,d)=1} \frac{a(s)}{s} S(-hd\bar{J}, \pm k; s) G(h, \mp k; s)$$

where $a(s)$ is a function of C^2 class with support $[S, 2S]$, $S \leq 2\sqrt{P}$ and satisfying $a^{(l)}(s) \ll S^{-l}$ for $l = 0, 1, 2$. The terms with $|k| \geq DSP^{-1/2}x^{3\epsilon} = K$, say, can be eliminated trivially: integrate by parts $l = [4\epsilon^{-1}]$ times with respect to ξ in $G(h, \pm k, s)$, getting

$$G(h, \pm k, s) = \left(\frac{-ds}{2\pi i k} \right)^l \int \frac{\partial^l}{\partial \xi^l} g(\xi^2 + s^2, h) e\left(\frac{\xi k}{ds} \right) d\xi \\ \ll \left(\frac{ds}{|k|\sqrt{P}} x^{2\epsilon} \right)^l \sqrt{P} \ll k^{-2} x^{-1}.$$

Therefore such terms contribute to $R(x,d,P)$ less than

$$\frac{\tau(d)}{d} \sum_{0 < |h| \leq H} \sum_{k \geq 1} \frac{1}{k^2 x} \sum_{0 < s \leq 2\sqrt{P}} 1 \ll \frac{P^{3/2}}{dx^2} x^\epsilon \ll \frac{x^{1-\epsilon}}{d}.$$

For $0 \leq |h| \leq H$, $0 < |k| \leq K$ and $S < s \leq 2S$ we trivially have

$$(15) \quad \frac{\partial^{l_1+l_2+l_3}}{\partial h^{l_1} \partial k^{l_2} \partial s^{l_3}} G(h, \mp k, s) \frac{a(s)}{s} \ll |h|^{-l_1} |k|^{-l_2} |s|^{-l_3} \frac{x^{1+12\epsilon}}{S\sqrt{P}}$$

for $0 \leq l_1, l_2, l_3 \leq 2$. This shows that Lemma 3 is applicable with

$$f(h,k,s) = \frac{S\sqrt{P}}{x^{1+13\epsilon}} \frac{a(s)}{s} G(h, \mp k, s) \text{ giving}$$

$$\left(\sum_{D < d \leq 2D} |R(x,d,P)| \right)^2 \ll x^{2-2\epsilon} + \frac{x^{2+40\epsilon}HK}{D^2P} \\ \times \sup_{1 \leq s \leq 2\sqrt{P}} \left\{ \frac{D^2(DS^2 + HK + HS^2)(DS^2 + HK + KS^2)}{S^2(DS^2 + HK)} + DS\sqrt{D(D+K)} \right\} \\ \ll x^{2-2\epsilon} + (D^2x + DP + DxP^{\frac{1}{2}})x^{48\epsilon}.$$

Therefore (10) holds if

(16)

$$D \leq x^{\frac{1}{2} - 25\epsilon}, \quad D \leq x^{2-10\epsilon} P^{-1} \quad \text{and} \quad D \leq x^{1-10\epsilon} P^{-\frac{1}{2}}.$$

This result can be compared with Hooley's $D = x^{1-\epsilon} P^{-3/4} \dots$

8. Evaluation of the main term.

For d square-free with $\omega(d) \neq 0$ consider

$$L(s, d) = \sum_{m=1}^{\infty} \frac{\omega(dm)}{\omega(d)} m^{-s}.$$

LEMMA 4. — *We have*

$$(17) \quad L(s, d) = \frac{\zeta(s)L(s, \chi_d)}{\zeta(2s)} \prod_{p|d} \left(1 + \frac{1}{p^s}\right)^{-1}.$$

Proof. — Follow the arguments of [3] on pp. 31-32 and equation (6.1).
Writing

$$C(m) \frac{\log m}{m} = \frac{1}{2\pi i} \int_{(\sigma)} R(s) m^{-s} ds, \quad \sigma > 0$$

by Mellin's inversion formula and partial integration two times

$$R(s) = \int C(\xi) \frac{\log \xi}{\xi} \xi^{s-1} d\xi \ll (|s|+1)^{-2} P^{\sigma-1} \log P.$$

Therefore

$$\begin{aligned} \sum_{m \equiv 0 \pmod{d}} \frac{\omega(m)}{m} c(m) \log m &= \frac{1}{2\pi i} \int_{(\sigma)} R(s) \frac{\omega(d)}{d^s} L(s, d) ds \\ &= R(1) \frac{\omega(d)}{d} \frac{L(1, \chi_d)}{\zeta(2)} \prod_{p|d} \left(1 + \frac{1}{p}\right)^{-1} + \frac{1}{2\pi i} \int_{\left(\frac{1}{2}\right)} R(s) \frac{\omega(d)}{d^s} L(s, d) ds \\ &= \frac{\omega(d)}{d} \prod_{p|d} \left(1 + \frac{1}{p}\right)^{-1} \frac{L(1, \chi_d)}{\zeta(2)} \int C(\xi) \frac{\log \xi}{\xi} d\xi + O\left(\frac{\tau^2(d)}{\sqrt{dP}} \log P\right). \end{aligned}$$

This yields

$$V(x, P) = \left(\sum_{d < D} \frac{\lambda_d}{d} \rho(d) \right) \frac{L(1, \chi_4)}{\zeta(2)} \int C(\xi) \frac{\log \xi}{\xi} d\xi + O\left(\sqrt{\frac{D}{P}} (\log x)^4\right)$$

where $\rho(d) = \omega(d) \prod_{p|d} \left(1 + \frac{1}{p}\right)^{-1}$. Now we specify λ_d to be those of the Rosser sieve giving (see [4])

$$\begin{aligned} \sum_{d < D} \lambda_d \frac{\rho(d)}{d} &= \prod_{p < D} \left(1 - \frac{\rho(p)}{p}\right) \left(2e^\gamma + O\left(\frac{1}{\log D}\right)\right) \\ &= \prod_{p < D} \left(1 - \frac{1}{p}\right) \frac{\zeta(2)}{L(1, \chi_4)} \left(2e^\gamma + O\left(\frac{1}{\log D}\right)\right) \\ &= \frac{2\zeta(2)}{L(1, \chi_4)} \frac{1}{\log D} \left(1 + O\left(\frac{1}{\log D}\right)\right) \end{aligned}$$

by the Mertens prime number theorem. Hence we conclude that

$$V(x, P) = \frac{2}{\log D} \int C(\xi) \frac{\log \xi}{\xi} d\xi \left(1 + O\left(\frac{1}{\log D}\right)\right).$$

We choose D equal to $x^{1-10\varepsilon} P^{-\frac{1}{2}}$ thus by (6) the total main term is equal to

$$\begin{aligned} X \sum_{0 \leq j \leq 1} V(x, P_j) &= 2(1 + O(\varepsilon)) X \int_x^{P_x} \frac{\text{Log } \xi}{\xi \text{Log}(x/\sqrt{\xi})} d\xi \\ &= 2(1 + O(\varepsilon)) X \int_1^\theta \frac{t dt}{1 - t/2} \text{Log } x \\ &= (1 + O(\varepsilon)) f(\theta) X \text{Log } x \end{aligned}$$

where $f(\theta) = 4(1 - \theta - 2 \text{Log}(2 - \theta))$ and is less than 1 for $\theta = 1.202\ 468\ 87$. The proof the Theorem follows from this and (3).

One may note that the truth of Selberg's eigenvalue conjecture leads to the lower bound $x^{\sqrt{3/2} - \varepsilon}$ for P_x .

BIBLIOGRAPHY

- [1] J.-M. DESHOUILERS and H. IWANIEC, Kloosterman sums and Fourier coefficients of cusp forms, *Inv. Math.* (to appear).

- [2] C. HOOLEY, On the greatest prime factor of a quadratic polynomial, *Acta Math.*, 117 (1967), 281-299.
- [3] C. HOOLEY, *Applications of sieve methods to the theory of numbers*, Cambridge Univ. Press, London, 1976.
- [4] H. IWANIEC, Rosser's sieve, *Acta Arith.*, 36 (1980), 171-202.
- [5] H. J. S. SMITH, Report on the theory of numbers, *Collected Mathematical Papers*, vol. I, reprinted, Chelsea, 1965.

Manuscrit reçu le 17 mars 1981
révisé le 15 octobre 1981.

J.-M. DESHOUILLERS,
Université de Bordeaux I
U.E.R. de Mathématiques
et d'Informatique
Laboratoire associé au CNRS n° 226
351, cours de la Libération
F - 33405 Talence Cedex.

H. IWANIEC,
Mathematics Institute
Polish Academy of Sciences
ul. Śniadeckich 8
PL - 00950 Warszawa.
