ACTA ARITHMETICA
106.2 (2003)

Primes in short arithmetic progressions

by

Jan-Christoph Puchta (Oxford)

The large sieve inequality in the form

$$\sum_{q\le Q} q \sum_{a=1}^{q}{}^{*} \Big| \sum_{\substack{n\le N\\ n\equiv a \,(\mathrm{mod}\, q)}} a_n - \frac{1}{q}\sum_{n\le N} a_n \Big|^2 < (N+Q^2)\sum_{n\le N}|a_n|^2$$

is essentially optimal. However, in several applications many of the a_n vanish, and one might expect better estimates then. In fact, such estimates were given by P. D. T. A. Elliott [1]. He showed the following estimate:

Theorem 1. *Let N and Q be integers, a_p be complex numbers for all primes $p \le N$. Then we have the estimate*

$$\sum_{q\le Q}(q-1)\sum_{(a,q)=1}\Big|\sum_{\substack{p\le N\\ p\equiv a \,(\mathrm{mod}\, q)}} a_p - \frac{q}{\varphi(q)}\sum_{p\le N} a_p\Big|^2 \ll_\varepsilon \left(\frac{N}{\log N} + Q^{54/11+\varepsilon}\right)\sum_{p\le N}|a_p|^2.$$

Under GRH, $Q^{54/11}$ may be replaced by $Q^{4+\varepsilon}$. In analogy to the large sieve, he conjectured that one may replace this term by $Q^{2+\varepsilon}$.

Using a completely different approach, Y. Motohashi [4] showed that

$$\sum_{q\le Q}\sum_{\chi \,(\mathrm{mod}\, q)}^{*} |\pi(x,\chi)|^2 \le \frac{(2+o(1))x^2}{\log x \log(x/Q^{1/2})} \tag{1}$$

for $x > Q^{5+\varepsilon}$, where $\pi(x,\chi) = \sum_{p\le x}\chi(p)$. He also conjectured that $Q^{5+\varepsilon}$ may be replaced by $Q^{2+\varepsilon}$.

Here we will combine the large sieve principle with Selberg's sieve to prove the conjecture of Elliott and give a version of (1) valid for $x > Q^{2+\varepsilon}$.

2000 *Mathematics Subject Classification*: 11N35, 11N13.
Key words and phrases: large sieve, Selberg's sieve.

I would like to thank D. R. Heath-Brown for his help with Proposition 9 which allowed me to reduce the exponent to $2+\varepsilon$, and the referee for pointing out some mistakes.

THEOREM 2. *Let N and Q be integers with $N > Q^{2+\varepsilon}$, a_p be complex numbers for any prime $p \le N$, and let $2 \le R \le \sqrt{N}$ be an integer. Then*

$$\sum_{q\le Q}\sideset{}{^*}\sum_{\chi \,(\mathrm{mod}\, q)}\Big|\sum_{p\le N} a_p\chi(p)\Big|^2 \ll_\varepsilon \frac{N}{\log N}\sum_{p\le N}|a_p|^2.$$

As this estimate is the analogue of the large sieve estimate, we can give analogues of Halász-type inequalities, too. As there are a variety of different large value estimates, the same is true for these bounds. However, since the optimal estimate depends on the particular application, we only mention the following:

THEOREM 3. *Let q be an integer. Let $\mathcal{C}$ be a set of characters $(\mathrm{mod}\, q)$, and a_p be complex numbers for any prime $p \le N$. Then for $k = 2, 3$ or, if q is cubefree, for any integer $k \ge 2$, we have the estimates*

$$\sum_{\chi\in\mathcal{C}}\Big|\sum_{p\le N} a_p\chi(p)\Big|^2 \le \left(\frac{N}{\log R} + c_{k,\varepsilon}N^{1-1/k}q^{(k+1)/(4k^2)+\varepsilon}|\mathcal{C}|R^{2/k}\right)\sum_{p\le N}|a_p|^2$$

and

$$\sum_{\chi\in\mathcal{C}}\Big|\sum_{p\le N} a_p\chi(p)\Big|^2 \le \left(\frac{N}{\log R} + R^2|\mathcal{C}|\sqrt{q}\,\log q\right)\sum_{p\le N}|a_p|^2.$$

If $\mathcal{C}$ is a set of characters to moduli $q \le Q$, the same bounds apply with q replaced by Q^2, where k can be chosen arbitrarily if all occurring values of q are cubefree, and $k = 2, 3$ otherwise.

From this we conclude immediately

COROLLARY 4. *For $x > Q^{2+\varepsilon}$ we have the estimate*

$$\sum_{q\le Q}\sideset{}{^*}\sum_{\chi \,(\mathrm{mod}\, q)}|\pi(x,\chi)|^2 \le C_\varepsilon\frac{x^2}{\log^2 x}.$$

Moreover, for $x > Q^{3+\varepsilon}$ this can be made completely explicit:

$$\sum_{q\le Q}\sideset{}{^*}\sum_{\chi \,(\mathrm{mod}\, q)}|\pi(x,\chi)|^2 \le \frac{(2+o(1))x^2}{\log x\log(x/Q^3)}.$$

We can also consider a single character:

COROLLARY 5. *Let χ be a complex character. Then*

$$|\pi(x,\chi)| \le \left(\left(\frac{1+\phi/\alpha}{2-2\phi/\alpha}\right)^{1/2} + o(1)\right)\frac{x}{\log x},$$

where $\alpha = \log x/\log q$ and $\phi = 1/4$ if q is cubefree, and $\phi = 1/3$ otherwise.

Note that this estimate is nontrivial if $x > q^{3/4}$ resp. $x > q$, depending on whether q is cubefree or not. With a little more work, we obtain the following statement.

COROLLARY 6. *Let D, x, Q be parameters with $x > Q^{1+\varepsilon}D^2$. Let N be the number of moduli $q \le Q$ such that there is some primitive character χ of order $d \le D$ and some dth root of unity ζ such that there is no prime $p \le x$ with $\chi(p) = \zeta$. Then $N \ll_\varepsilon D$.*

This was proven by Elliott with $D = 3$ under the condition $x > Q^{54/11+\varepsilon}$.

We begin the proof of our theorems with the following two sieve principles.

LEMMA 7 (Bombieri). *Let $V, (\cdot,\cdot)$ be an inner product space, $v_i \in V$. Then for any $\Phi \in V$ we have*

$$\sum_i |(\Phi, v_i)|^2 \le \|\Phi\|^2 \max_i \sum_j |(v_i, v_j)|.$$

This is Lemma 1.5 of [3].

LEMMA 8 (Selberg). *Let R, N be integers such that $R^2 < N$. Then there is a function g which has the following properties*:

1. $g(1) = 1$, $|g(n)| \le 1$ *for* $n \le R$, $g(n) = 0$ *for* $n > R$.
2. $\sum_{n\le N}((1 * g)(n))^2 \le N/\log R + R^2$.

This is the usual formulation of Selberg's sieve when used to count the set of primes $\le N$ (see e.g. [2, Chapter 3, especially Theorem 3.3]). In what follows, we will denote by g the function given by Lemma 8 and set $f = (1 * g)^2$. We will have to bound character sums involving f; these computations are summarized in the following proposition.

PROPOSITION 9. *Let $\chi \pmod q$ be a character, R, N, f and g as in Lemma* 8, *and define $S = \sum_{n\le N} f(n)\chi(n)$.*

1. *If χ is principal, we have $|S| < N/\log R + R^2$.*

2. *Assume that χ is nonprincipal. Then for any fixed A we have the estimate*

$$\sum_{\nu=1}^{\infty} f(\nu)\chi(\nu)e^{-\log^2(\nu/N)} \ll_{\varepsilon,A} R^2 q^{1/2}\left(\frac{N}{R^2 q}\right)^{-A}.$$

3. *If χ is nonprincipal, we have the bounds $|S| \le R^2\sqrt{q}\log q$ and $|S| \le c_{k,\varepsilon}R^{2/k}N^{1-1/k}q^{(k+1)/(4k^2)+\varepsilon}$ for $k = 2, 3$, or, if q is cubefree, for $k \ge 2$ arbitrary.*

Proof. The first assertion is already contained in Lemma 8.

Assume now that χ is nonprincipal. Then

$$\begin{aligned}\Big|\sum_{n\le N} f(n)\chi(n)\Big| &= \Big|\sum_{n\le N}\Big(\sum_{d|n} g(d)\Big)^2\chi(n)\Big| \\ &= \Big|\sum_{d_1,d_2\le R} g(d_1)g(d_2)\chi([d_1,d_2]) \sum_{n\le N/[d_1,d_2]}\chi(n)\Big| \\ &\le \sum_{d_1,d_2\le N} |g(d_1)g(d_2)|\cdot\Big|\sum_{n\le N/[d_1,d_2]}\chi(n)\Big| \\ &\le \sum_{d_1,d_2\le R}\Big|\sum_{n\le N/[d_1,d_2]}\chi(n)\Big|.\end{aligned}$$

The inner sum can be estimated using either the Pólya–Vinogradov inequality or Burgess estimates, leading to $|S| \le R^2\sqrt{q}\log q$, resp. $|S| \le c_{k,\varepsilon}R^{2/k}N^{1-1/k}q^{(k+1)/(4k^2)+\varepsilon}$; thus we obtain the third statement.

To prove the second statement, we begin as above to obtain the inequality

$$\Big|\sum_{n=1}^{\infty} f(n)\chi(n)e^{-\log^2(n/N)}\Big| \le \sum_{d_1,d_2\le R}\Big|\sum_{n=1}^{\infty}\chi(n)e^{-\log^2([d_1,d_2]n/N)}\Big|.$$

Write $d = [d_1,d_2]$. Using the Mellin transform

$$\frac{1}{2\sqrt{\pi}\,i}\int\limits_{(2)} x^{-s}e^{s^2/4}\,ds = e^{-\log^2 x},$$

the inner sum can be expressed as

$$\sum_{n=1}^{\infty}\chi(n)e^{-\log^2(dn/N)} = \frac{1}{2\sqrt{\pi}\,i}\int\limits_{(2)} L(s,\chi)e^{s^2/4}(N/d)^s\,ds.$$

Now we shift the path of integration to the line $\Re s = -A$ with $A > 0$. Denote by χ_1 the primitive character inducing χ. Then

$$L(s,\chi) = \prod_{p|q_2}(1-\chi_1(p)p^{-s})L(s,\chi_1).$$

For $A > 2$, the first factor is $\ll q_2^A$, whereas using the functional equation the L-series can be estimated to be $\ll (q_1(|t|+2))^{A+1/2}$, hence the right hand side is

$$\ll_A q^{1/2}\left(\frac{N}{dq}\right)^{-A} \le q^{1/2}\left(\frac{N}{R^2q}\right)^{-A}.$$

Hence the whole sum can be bounded by $c(A)R^2q^{1/2}(N/(R^2q))^{-A}$.

To prove Theorem 2, we follow the lines of the proof of the large sieve resp. the Halász inequality; however, we apply Lemma 7 to a different euclidean space. Consider the subspace $V < l^\infty$ consisting of all bounded

sequences (a_n) such that $a_n = 0$ whenever $f(n) = 0$, where f is defined as in Lemma 8. On this space define a product as

$$\langle (a_n), (b_n) \rangle := \sum_{n=1}^{\infty} f(n) e^{-\log^2(n/N)} a_n \overline{b}_n.$$

Now we apply Lemma 7 to this space and the set of vectors $\Phi = (\widehat{a}_n)$, where $\widehat{a}_p = a_p e^{\log^2(p/N)}$, for prime numbers p in the range $R^2 < p \le N$, and $\widehat{a}_n = 0$ otherwise, and $v_i = (\widehat{\chi}(n))$, where similarly $\widehat{\chi}(n) = \overline{\chi(n)}$, if $f(n) \ne 0$, and 0 otherwise. Now the inequality reads as

$$\begin{aligned}\sum_{q \le Q} \sum_{\chi \,(\mathrm{mod}\, q)} \Big| \sum_{R^2 < p \le N} a_p \chi(p) \Big|^2 \\ \le \max_{\chi} \Big(\sum_{n=1}^{\infty} f(n) e^{-\log^2(n/N)} + \sum_{\chi' \ne \chi} \Big| \sum_{n \le N} f(n) e^{-\log^2(n/N)} \chi\overline{\chi'}(n) \Big| \Big) \\ \times \sum_{p \le N} |a_p|^2 e^{2\log^2(p/N)}\end{aligned}$$

where the maximum is taken over all characters with moduli at most Q. From Lemma 8 it follows that the first term inside the brackets is $\ll N/\log R$ provided that $R < N^{1/3}$, say. For the second term, let χ be a character $(\mathrm{mod}\, q)$ and χ' a character $(\mathrm{mod}\, q')$. Then $\chi\overline{\chi'}$ is a character $(\mathrm{mod}\, [q, q'])$. By Proposition 9, each term in the outer sum can be bounded by $c(A) R^2 [q, q']^{1/2} \times (N/(R^2[q, q']))^{-A}$, hence the whole sum is $\le c(A) Q^3 R^2 (N/(R^2 Q^2))^{-A}$. Since by assumption $N > Q^{2+\varepsilon}$, we can choose $R = Q^{\varepsilon/4}$, $A = 6/\varepsilon + 1$ to bound this by some constant depending only on ε. Thus we get the estimate

$$\sum_{q \le Q} \sum_{\chi \,(\mathrm{mod}\, q)} \Big| \sum_{R^2 < p \le N} a_p \chi(p) \Big|^2 \ll \Big(\frac{N}{\varepsilon \log N} + C_\varepsilon \Big) \sum_{p \le N} |a_p|^2.$$

The range $n \le R^2$ can be estimated using the usual large sieve inequality, which gives $(R^2 + Q^2) \sum_{p \le N} |a_p|^2$, which is negligible. Hence Theorem 2 is proven.

The proof of Theorem 3 is similar, but simpler. First, assume that all characters in $\mathcal{C}$ are characters to a single modulus q. We consider the vector space $V < \mathbb{C}^N$ consisting of the sequences $(a_n)_{n=1}^N$ with $a_n = 0$ for all n with $f(n) = 0$ and the scalar product $\langle (a_n), (b_n) \rangle := \sum_{n \le N} f(n) a_n \overline{b}_n$. Applying Lemma 7 as above, we obtain the estimate

$$\sum_{\chi \in \mathcal{C}} \Big| \sum_{p \le N} a_p \chi(p) \Big|^2 \le \Big(\frac{N}{\log R} + R^2 + (|\mathcal{C}| - 1) \Delta(R, N, q) \Big) \sum_{R \le p \le N} |a_p|^2$$

where $\Delta(R,N,q)$ is the bound obtained by Proposition 9, i.e. $\Delta(R,N,q) \leq R^2\sqrt{q}\log q$, resp. $\Delta(R,N,q) < c_{k,\varepsilon}q^{(k+1)/(4k^2)+\varepsilon}N^{1-1/k}R^{2/k}$. The term R^2 can be neglected in comparison with $\Delta(R,N,q)$. This is obvious in the first case. In the second case, we may assume that $\Delta(R,N,q) < N$, since otherwise Theorem 3 is an immediate consequence of the Cauchy–Schwarz inequality. This implies $R < N^{1/2}q^{-(k+1)/(2k)}$, which in turn implies

$$R^2 < N^{1-1/k}q^{-1-1/k} < \Delta(R,N,q).$$

Hence we obtain Theorem 3 for sets of characters belonging to a single modulus.

The proof for the case that the characters belong to different moduli is similar; note that $[q_1,q_2]$ is cubefree if both q_1 and q_2 are cubefree.

In the range $Q^{2+\varepsilon} \leq x < Q^{3+\varepsilon}$, Corollary 4 follows from Theorem 2 by choosing $a_p = 1$ for all prime numbers $p \leq N$, whereas in the range $x > Q^{3+\varepsilon}$ it follows from Theorem 3. Similarly we obtain Corollary 5 from Theorem 3. We choose $\mathcal{C} = \{\chi_0, \chi, \overline{\chi}\}$ to obtain the estimate

$$|\pi(x)|^2 + 2|\pi(x,\chi)|^2 \leq \frac{x}{\log(c_{k,\varepsilon}x^{1/2}q^{(k+1)/(8k)+\varepsilon})}\,\pi(x)$$

and choosing either $k = 3$ or $k \nearrow \infty$ we obtain the result by solving for $|\pi(x,\chi)|$.

To prove Corollary 6, let $\mathcal{P}$ be the set of prime numbers p, such that there is some character χ of order d as described in the corollary. For every such p, choose such a character χ_1 together with all its powers, and denote the set of all these characters with $\mathcal{C}$. Let ζ be a dth root of unity. We have

$$\sum_{\substack{\chi^d=\chi_0\\ \chi\neq\chi_0}} |\pi(x,\chi)|^2 = d\sum_{a=1}^{d}\left|\#\{p\leq x \mid \chi_1(p)=\zeta^a\} - \frac{1}{d}\pi(x,\chi_0)\right|^2.$$

Since by assumption, one of the terms on the right hand side is large, the right hand side is

$$\gg \frac{x^2}{d\log^2 x} \geq \frac{x}{D\log^2 x}.$$

Now we have $|\mathcal{C}| \leq D\cdot|\mathcal{P}|$; thus we get

$$|\mathcal{P}|\,\frac{x^2}{D\log^2 x} \ll \frac{x^2}{\log x\log R} + xDR^2|\mathcal{P}|Q\log Q.$$

If $D^2Q\log Q < x^{1-\varepsilon}$, we can choose $R = x^{\varepsilon/4}$, and the second term on the right hand side is still of lesser order than the left hand side. With this choice the inequality can be simplified to $|\mathcal{P}| \ll_\varepsilon D$.

References

[1] P. D. T. A. Elliott, *Subsequences of primes in residue classes to prime moduli*, in: Studies in Pure Mathematics to the Memory of P. Turán, P. Erdős (ed.), Akadémiai Kiadó, Budapest, 1983, 157–164.

[2] H. Halberstam and H.-E. Richert, *Sieve Methods*, London Math. Soc. Monographs 4, Academic Press, 1974.

[3] H. L. Montgomery, *Topics in Multiplicative Number Theory*, Lecture Notes in Math. 227, Springer, 1971.

[4] Y. Motohashi, *Large sieve extensions of the Brun–Titchmarsh theorem*, in: Studies in Pure Mathematics to the Memory of P. Turán, P. Erdős (ed.), Akadémiai Kiadó, Budapest, 1983, 507–515.

Mathematical Institute
University of Oxford
24-29 St. Giles' Street
Oxford, OX1 3LB, U.K.
E-mail: puchta@maths.ox.ac.uk

Current address:
Mathematisches Institut
Eckerstr. 1
79111 Freiburg, Germany
E-mail: jcp@arcade.mathematik.uni-freiburg.de

Received on 13.6.2001
and in revised form on 22.4.2002 (4051)