NONNEGATIVE MULTIPLICATIVE FUNCTIONS ON SIFTED SETS, AND THE SQUARE ROOTS OF -1 MODULO SHIFTED PRIMES

PAUL POLLACK

Abstract. An oft-cited result of Peter Shiu bounds the mean value of a nonnegative multiplicative function over a coprime arithmetic progression. We prove a variant where the arithmetic progression is replaced by a sifted set. As an application, we show that the normalized square roots of $-1 \pmod m$ are equidistributed (mod 1) as m runs through the shifted primes $q-1$.

1. Introduction

Many problems in elementary and analytic number theory require estimates for mean values of arithmetic functions. One of our most useful tools for obtaining upper bound estimates is the following theorem of Peter Shiu [Shi80], which bounds the mean value of a nonnegative-valued multiplicative function over a coprime arithmetic progression.

Let $\mathscr{M}$ be the collection of nonnegative-valued multiplicative functions f satisfying the following two conditions:

(I) There is a constant $A_1 > 1$ such that $f(p^k) \le A_1^k$ for all prime powers p^k.

(II) For every $\epsilon > 0$, there is a constant $A_2(\epsilon) > 0$ such that

$$f(n) \le A_2(\epsilon) n^{\epsilon} \qquad \text{for all } n \ge 1.$$

Theorem A (Brun–Titchmarsh for multiplicative functions, [Shi80])**.** *Let $f \in \mathscr{M}$, $0 < \alpha, \beta < \frac{1}{2}$, and let a, k be integers satisfying $0 \le a < k$ and $\gcd(a,k) = 1$. Then for all sufficiently large x, we have that*

$$\sum_{\substack{x-y<n\le x \\ n \equiv a \pmod k}} f(n) \ll \frac{y}{\varphi(k)} \frac{1}{\log x} \exp\left(\sum_{\substack{p \le x \\ p \nmid k}} \frac{f(p)}{p}\right), \tag{1.1}$$

whenever a, k, y satisfy

$$k < y^{1-\alpha}, \quad x^{\beta} < y \le x.$$

Here the implied constant in (1.1), as well as the thresshhold for "sufficiently large", depends only on α, β, the constant A_1 in (I) above, and the function $A_2(\epsilon)$ in (II).

Theorem A has proved highly influential, with AMS Mathematical Reviews recording more than 70 citations to [Shi80] so far. No doubt this is due to its impressive generality and the ease with which it can be "plugged in" as an auxiliary tool in number-theoretic investigations. Shiu's work is based on methods described by Erdős [Erd52] and Wolke [Wol71]; work in a similar direction had also been carried out by Vinogradov and Linnik [VL57], Barban [Bar64a], and Barban and Vehov [BV69].

In this paper, we put forward a variant of Theorem A where the role of the coprime progression a mod k is taken by a sifted set.

2010 *Mathematics Subject Classification.* 11N37 (primary), 11N36, 11J71 (secondary).

Theorem 1.1 (Brun's upper bound sieve for multiplicative functions)**.** *Let $f \in \mathscr{M}$, let $0 < \beta < \frac{1}{2}$, and let k be a nonnegative integer. For each prime $p \le x$, let $\mathscr{E}_p$ be a union of $\nu(p)$ nonzero residue classes modulo p, where we suppose that each $\nu(p) \le k$. Let*

$$\mathscr{S} = \bigcap_{p \le x} \mathscr{E}_p^{\mathsf{c}}.$$

(In words, $\mathscr{S}$ is the set of all integers n not belonging to any $\mathscr{E}_p$.) If x is sufficiently large, then

$$\sum_{\substack{x-y<n\le x\\ n\in\mathscr{S}}} f(n) \ll \frac{y}{\log x}\exp\left(\sum_{p\le x}\frac{f(p)-\nu(p)}{p}\right) \tag{1.2}$$

for all y satisfying

$$x^{\beta} < y \le x.$$

Here the implied constant in (1.2)*, as well as the thresshhold for "sufficiently large", depends only on β, k, the constant A_1 in* (I) *above, and the function $A_2(\epsilon)$ in* (II)*.*

Remarks.

(i) Theorem 1.1, while obviously a close relative of Theorem A, does not obviously imply it (nor vice versa).

(ii) It may initially seem strange that we require $\mathscr{E}_p$ to only contain nonzero residue classes. This does not entail any loss of generality, since we can effectively remove n not coprime to a given P by replacing the function $f(n)$ with $\mathbb{1}_{\gcd(n,P)} \cdot f(n)$.

(iii) Keeping the last remark in mind, one easily deduces from Theorem 1.1 that the number of $n \le x$ that avoid any prescribed $\nu(p) \le k$ residue classes modulo p, for each prime $p \le x$, is

$$\ll_k x \exp\left(-\sum_{p\le x}\frac{\nu(p)}{p}\right).$$

In this way, Theorem 1.1 generalizes Brun's upper bound sieve.

As a first example of Theorem 1.1, if we remove a single nonzero residue class modulo p for each prime $p \le x^{1/2}$, then the sum of the divisor function $d(n)$ over the remaining $n \in [1, x]$ is $O(x)$, where the implied constant is absolute.

Another immediate application of Theorem 1.1 is an upper bound for the mean value of $f(n)$ with n restricted to shifted twin primes.

Corollary 1.2. *Let f be a function belonging to $\mathscr{M}$. Let $0 < \beta < 1/2$. For all $x \ge 3$ and $y \in (x^{\beta}, x]$,*

$$\sum_{\substack{x-y<q\le x\\ q-2,\ q \text{ prime}}} f(q-1) \ll_{A_1,A_2(\epsilon),\beta} \frac{x}{(\log x)^2}\exp\left(\sum_{p\le x}\frac{f(p)-1}{p}\right),$$

Taking $f(n) = t^{\omega(n)}$ and $y = x$ in Corollary 1.2, we deduce that for each fixed t_0, and all $x \ge 3$,

$$\sum_{\substack{q\le x\\ q-2,\ q \text{ prime}}} t^{\omega(q-1)} \ll_{t_0} \frac{x}{(\log x)^2}(\log x)^{t-1},$$

uniformly for $0 < t \le t_0$. One can now extract information on the distribution of $\omega(q-1)$ by varying t. For example, mimicking the proof of Theorem 010 in [HT88] yields: *Uniformly for*

$0 \le \psi \le (\log\log x)^{1/6}$, *the number of prime pairs* $q-2, q$ *in* $[1,x]$ *with* $|\omega(q-1) - \log\log x| > \psi\sqrt{\log\log x}$ *is* $O(x(\log x)^{-2}\exp(-\psi^2/2))$. This last statement is a strengthened form of a theorem of Barban [Bar64b].

We now describe our original motivation for proving Theorem 1.1. The application is a riff on two theorems of Hooley.

An infamous conjecture of Landau (one of his "four unattackable problems") predicts that there are infinitely many primes of the form x^2+1. This is still open, but the analogous problem for primes of the form x^2+y^2+1 was settled by Hooley in 1957 [Hoo57]. Put

$$r(n) = \#\{(x,y) \in \mathbb{Z}^2 : x^2+y^2 = n\}.$$

Hooley shows not only that $r(q-1) > 0$ for infinitely many primes q, but that $r(q-1)$ has a positive average value.

Theorem B. *For a certain positive constant* K*, we have*

$$\sum_{\substack{q \le x \\ q \text{ prime}}} r(q-1) \sim K\frac{x}{\log x} \qquad (x \to \infty).$$

Actually, Hooley's work was conditional on GRH, but the discovery of the Bombieri–Vinogradov theorem allowed for this dependence to be removed with minimal changes to Hooley's argument. See [EH66]. (In the intervening years, Linnik [Lin60] gave an alternative proof of Theorem B.)

The following variant of Hooley's result was shown by Kátai in 1968 [Kát68] (see also [Kát70]).

Theorem B′. *For a certain positive constant* K'*, we have*

$$\sum_{\substack{q \le x \\ q \text{ prime} \\ q-1 \text{ squarefree}}} r(q-1) \sim K'\frac{x}{\log x} \qquad (x \to \infty).$$

From elementary number theory, we know that for squarefree n, the number of square roots of -1 modulo n is precisely $\frac{r(n)}{4}$. In particular, Theorem B′ implies that -1 is a square modulo $q-1$ for infinitely many primes q.

We now bring in the second theorem of Hooley. First, a definition: If $P(T)$ is a polynomial with integer coefficients, and k is any positive integer, then a *normalized root of* P*, modulo* k, is a rational number of the form ϖ/k, where

$$f(\varpi) \equiv 0 \pmod{k}, \qquad 0 \le \varpi < k.$$

In 1964, Hooley proved the following equidistribution theorem for normalized roots [Hoo64].

Theorem C. *Let* $P(T) \in \mathbb{Z}[T]$ *be an irreducible polynomial of degree at least* 2*. For each positive integer* k*, list the normalized roots of* P *modulo* k *(in any order), and then concatenate the lists sequentially for* $k = 1, 2, 3, \dots$*. The resulting sequence is uniformly distributed in* $[0,1)$*.*

After this set-up, the reader can perhaps guess where we are headed. We prove that when $P(T) = T^2+1$, the conclusion of Hooley's Theorem C holds with the moduli restricted to the shifted primes $q-1$.

Theorem 1.3. *For each prime q, list the normalized square roots of -1 modulo $q-1$, and then concatenate the lists successively for $q = 2, 3, 5, 7, \dots$. The resulting sequence is uniformly distributed in* $[0, 1)$.

Remark. By a deeper and very different method, Duke, Friedlander, and Iwaniec [DFI95] have shown that Theorem 1.3 holds with the moduli running over primes q rather than shifted primes $q-1$.

In both Corollary 1.2 and Theorem 1.3, (with some effort) we could have applied the Nair–Tenenbaum generalization of Shiu's theorem [NT98] in place of Theorem 1.1; this is because of the simple structure of the sieved-out residue classes in these examples. (For each prime p, the classes being sieved out can be described as the mod p roots of a polynomial of small height.) Nevertheless, we think that Theorem 1.1 (which, in full generality, does not follow trivially from [NT98]), and the mentioned applications, are worthy of attention in themselves. We are also hopeful that future researchers will find our description of the EBLSVVW[1] method useful; we have formulated Proposition 2.1 with this ambition in mind.

Notation. Most of our notation is standard. A possible exeption is our use of $P^+(n)$ and $P^-(n)$ for the largest and smallest prime factors of n (respectively); we adopt the convention that $P^+(1) = 1$ while $P^-(1) = \infty$. The notation $d \parallel n$ indicates that d is a unitary divisor of n; that is, $d \mid n$ and $\gcd(d, n/d) = 1$. We let $\mathbb{1}$ denote the function that is identically 1, and we use $\mathbb{1}_C$ for the characteristic function of a property or set C. We write $e(x)$ for $e^{2\pi i x}$. We reserve the letter p for prime numbers.

2. Proof of Theorem 1.1

2.1. **Generalities.** Let $f \in \mathscr{M}$, and let A_1 and $A_2(\epsilon)$ be as in (I) and (II). Let $x \geq 3$, and let $\theta \in (0, 1)$. For each integer $n \in [1, x]$, we may write

$$n = p_1 \dots p_j p_{j+1} \cdots p_J,$$

where

$$p_1 \leq p_2 \leq \cdots \leq p_J,$$

and where j is chosen as the largest index for which

$$p_1 \cdots p_j \leq x^\theta \quad \text{and} \quad p_1 \cdots p_j \parallel n.$$

We let

$$d := p_1 \cdots p_j,$$

and we refer to d as the *canonical unitary prefix divisor* of n.

We make the assumption that n has no proper prime power divisor in the interval $(x^{\theta/2}, x^\theta]$ and that

$$\Omega(n/d) \geq 2\theta^{-1}.$$

(In our eventual application, we will be able to handle the excluded values of n by a separate argument.)

Since $p_{j+1}^{\Omega(n/d)} \leq p_{j+1} \cdots p_J \leq n \leq x$, we see that

$$p_{j+1} \leq x^{1/\Omega(n/d)} \leq x^{\theta/2}. \tag{2.1}$$

[1]Erdős–Barban–Linnik–Shiu–Vehov–Vinogradov–Wolke

Let t be the largest positive integer for which

$$p_{j+1} = p_{j+2} = \cdots = p_{j+t}.$$

Then $p_1 \cdots p_{j+t}$ is a unitary divisor of n, so that the choice of j forces $p_1 \cdots p_{j+t} > x^\theta$. Thus,

$$d = p_1 \cdots p_j > x^\theta / p_{j+1}^t.$$

If $p_{j+1}^t > x^{\theta/2}$, then $t > 1$, by (2.1); moreover, some power of p_{j+1} is then a proper prime power divisor of n belonging to $(x^{\theta/2}, x^\theta]$, contradicting our assumptions on n. So $p_{j+1}^t \le x^{\theta/2}$, and, from the last displayed equation,

$$d > x^{\theta/2}.$$

From (2.1), there is a (unique) integer $r \ge 2$ satisfying

$$x^{\theta/(r+1)} < p_j \le x^{\theta/r}.$$

Since

$$x \ge p_{j+1}^{\Omega(n/d)} \ge x^{\Omega(n/d)\cdot\theta/(r+1)},$$

we have $\Omega(n/d) \le (r+1)\cdot\theta^{-1}$, so that

$$\begin{aligned} f(n) &= f(p_1 \cdots p_j) f(p_{j+1} \cdots p_J) \\ &\le f(p_1 \cdots p_j) A_1^{\Omega(n/d)} \le f(d) A_1^{(r+1)\theta^{-1}}. \end{aligned}$$

We collect the salient results in the following proposition.

Proposition 2.1. *Let $f \in \mathscr{M}$, and let $\theta \in (0,1)$. Let $x \ge 3$. Let n be an integer in $[1,x]$, and assume n does not have a proper prime power divisor in $(x^{\theta/2}, x^\theta]$. Let d be the canonical unitary prefix divisor of n. If $\Omega(n/d) \ge 2/\theta$, then*

$$x^{\theta/2} < d \le x^\theta, \quad x^{\theta/(r+1)} < P^+(d) \le x^{\theta/r} \quad \text{for some integer } r \ge 2, \tag{2.2}$$

and

$$f(n) \le \exp(O(r)) f(d).$$

Here the implied constants depend only on θ, A_1, and $A_2(\epsilon)$.

2.2. **Completion of the proof of Theorem 1.1.** Fix $\theta = \beta/4$. We put those $n \in \mathscr{S}$ belonging to $(x-y, x]$ into the following three (possibly overlapping) categories. For each category we then bound the corresponding contribution to $\sum_{n \in \mathscr{S} \cap (x-y,x]} f(n)$.

(1) Arithmetically atypical n: Here we include all n with a proper prime power divisor in $(x^{\theta/2}, x^\theta]$. We also put in this category all n which have a $\log(x^\theta)$-smooth divisor in the interval $(x^{\theta/2}, x^\theta]$.
(2) Those n for which $\Omega(n/d) < 2/\theta$, where (as above) d is the canonical unitary prefix divisor.
(3) All remaining $n \in \mathscr{S}$.

We handle category (1) using the crude pointwise bound $f(n) \ll_\epsilon n^\epsilon$. The number of $n \in (x-y, x]$ divisible by a proper prime power $p^k \in (x^{\theta/2}, x^\theta]$ does not exceed

$$\sum_{\substack{x^{\theta/2} < p^k \le x^\theta \\ k \ge 2}} \left(\frac{y}{p^k} + 1\right) \le x^\theta + y \sum_{\substack{m > x^{\theta/2} \\ m \text{ squarefull}}} \frac{1}{m} \ll x^\theta + y x^{-\theta/4} \ll y x^{-\theta/4}.$$

Similarly, the number of $n \in (x-y, x]$ possessing a $\log(x^\theta)$-smooth divisor $e \in (x^{\theta/2}, x^\theta]$ is at most

$$x^\theta + y \sum_{\substack{x^{\theta/2} < e \le x^\theta \\ P^+(e) \le \log(x^\theta)}} \frac{1}{e} \le x^\theta + yx^{-\theta/2} \sum_{\substack{e \le x^\theta \\ P^+(e) \le \log(x^\theta)}} 1$$
$$\ll x^\theta + yx^{-\theta/3} \ll yx^{-\theta/3}.$$

(To justify passing from the first to the second line, we use that the count $\log T$-smooth numbers up to T is $T^{o(1)}$, as $T \to \infty$; see, e.g., [MV07, Corollary 7.9, p. 209].) Since $f(n) \ll x^{\theta/8}$ (say), it follows that the contribution to $\sum_{n \in \mathscr{S} \cap (x-y,x]} f(n)$ from arithmetically inconvenient n is

$$\ll (yx^{-\theta/4} + yx^{-\theta/3})x^{\theta/8} \ll yx^{-\theta/8}.$$

This is negligible, since the our target upper bound — the right-hand side of (1.2) — is $\gg y(\log x)^{-k-1}$.

For n falling into category (2), we have that $n = dp_{j+1} \cdots p_J$, where $J - j < 2\theta^{-1}$. If all of $p_{j+1}, p_{j+2}, \dots, p_J$ are bounded by $x^{\theta^2/2}$, then $n \le d(x^{\theta^2/2})^{2/\theta} \le x^{2\theta}$. Since $f(n) \ll x^\theta$, these n contribute $\ll x^{3\theta} \ll yx^{-\theta}$, which is once again negligible. So we may assume that $p_{j+t} > x^{\theta^2/2}$ for some positive integer $t \le J$. If t is chosen minimally, then putting

$$d' = p_1 p_2 \cdots p_{j+t-1},$$

we see that d' is a unitary divisor of n, that

$$d' \le x^{2\theta},$$

that

$$P^-(n/d') > x^{\theta^2/2},$$

and that

$$f(n) = f(d')f(n/d')$$
$$\le f(d')A_1^{2\theta^{-1}} \ll f(d').$$

Thus, these n make a contribution that is

$$\ll \sum_{d' \le x^{2\theta}} f(d') \sum_{\substack{(x-y)/d' < m \le x/d' \\ md' \in \mathscr{S} \\ P^-(m) > x^{\theta^2/2}}} 1. \tag{2.3}$$

We use Brun's sieve to bound the sum on m. Let p be a prime not exceeding $x^{\theta^2/2}$. The subscripted conditions imply that m avoids the residue class of 0 modulo p and, if $p \nmid d$, also an additional $\nu(p)$ residue classes modulo p. Since $y/d' \ge x^\beta/x^{2\theta} \ge x^{\theta^2/2}$, Brun's sieve bounds the sum as

$$\ll \frac{y}{d'} \prod_{\substack{p \le x^{\theta^2/2} \\ p \nmid d'}} \left(1 - \frac{\nu(p)+1}{p}\right) \prod_{\substack{p \le x^{\theta^2/2} \\ p \mid d'}} \left(1 - \frac{1}{p}\right)$$
$$\ll \frac{y}{d' \log x} \prod_{\substack{p \le x^{\theta^2/2} \\ p \nmid d'}} \left(1 - \frac{\nu(p)}{p}\right).$$

We can extend the product over all primes $p \le x$ without changing the order of magnitude. We conclude that

$$\sum_{\substack{\frac{x-y}{d'} < m \le \frac{x}{d'} \\ md' \in \mathscr{S} \\ P^-(m) > x^{\theta^2/2}}} 1 \ll \left(\frac{y}{d' \log x} \prod_{p \le x} \left(1 - \frac{\nu(p)}{p}\right)\right) \prod_{p \mid d'} \left(1 - \frac{\nu(p)}{p}\right)^{-1}$$

$$\ll \left(\frac{y}{d' \log x} \prod_{p \le x} \left(1 - \frac{\nu(p)}{p}\right)\right) g(d'),$$

where

$$g(d') := \prod_{p \mid d'} \left(1 - \frac{\min\{k, p-1\}}{p}\right)^{-1}.$$

Inserting these bounds back into (2.3), we find that the remaining n in category (2) contribute

$$\ll \left(\frac{y}{\log x} \prod_{p \le x} \left(1 - \frac{\nu(p)}{p}\right)\right) \sum_{d' \le x^{2\theta}} \frac{f(d')g(d')}{d'}$$

$$\ll \left(\frac{y}{\log x} \prod_{p \le x} \left(1 - \frac{\nu(p)}{p}\right)\right) \prod_{p \le x} \left(1 + \frac{f(p)g(p)}{p} + \frac{f(p^2)g(p^2)}{p^2} + \dots\right).$$

Now $f(p)g(p)/p = f(p)/p + O(1/p^2)$, while $\sum_{p \le x} \sum_{k \ge 2} f(p^k)g(p^k)/p^k \ll 1$. It follows that the final displayed product on p is $\ll \exp(\sum_{p \le x} f(p)/p)$, leading to an upper bound for the entire expression of

$$\ll \left(\frac{y}{\log x} \prod_{p \le x} \left(1 - \frac{\nu(p)}{p}\right)\right) \exp\left(\sum_{p \le x} \frac{f(p)}{p}\right),$$

which in turn is

$$\ll \frac{y}{\log x} \exp\left(\sum_{p \le x} \frac{f(p) - \nu(p)}{p}\right).$$

This expression coincides with that on the right-hand side of (1.2), and so the contribution from the n in category (2) is acceptable.

Finally, we turn to category (3). We subdivide the n in that category according to the value of r for which (2.2) holds. Since we are assuming that n does not have a $\log(x^\theta)$-smooth divisor in $(x^{\theta/2}, x^\theta]$, we only consider r for which $x^{\theta/r} > \log(x^\theta)$, so that

$$2 \le r < \frac{\log(x^\theta)}{\log\log(x^\theta)}. \tag{2.4}$$

By Proposition 2.1, for each fixed r in the range (2.4), the corresponding n make a contribution that is

$$\ll \exp(O(r)) \sum_{\substack{x^{\theta/2} < d \le x^\theta \\ P^+(d) \le x^{\theta/r}}} f(d) \sum_{\substack{(x-y)/d < m \le x/d \\ P^+(m) > x^{\theta/(r+1)} \\ md \in \mathscr{S}}} 1.$$

In this case, Brun's sieve gives that the sum on m is

$$\ll \frac{y}{d} \prod_{\substack{p \le x^{\theta/(r+1)} \\ p \nmid d}} \left(1 - \frac{\nu(p)+1}{p}\right) \prod_{\substack{p \le x^{\theta/(r+1)} \\ p \mid d}} \left(1 - \frac{1}{p}\right) \ll (r\theta^{-1})^{k+1} \left(\frac{y}{\log x} \prod_{p \le x} \left(1 - \frac{\nu(p)}{p}\right)\right) \frac{g(d)}{d},$$

where g has the same meaning as above. Since $(r\theta^{-1})^{k+1} = \exp(O(r))$, we see that these n contribute

$$\ll \exp(O(r)) \left(\frac{y}{\log x} \prod_{p \le x} \left(1 - \frac{\nu(p)}{p}\right)\right) \sum_{\substack{x^{\theta/2} < d \le x^{\theta} \\ P^+(d) \le x^{\theta/r}}} \frac{f(d)g(d)}{d}. \tag{2.5}$$

The sum on d is estimated by the following special case of [Shi80, Lemma 4].

Lemma 2.2. *Let $F \in \mathscr{M}$. Then, for all sufficiently large Z,*

$$\sum_{\substack{n \ge Z^{1/2} \\ P^+(n) \le Z^{1/R}}} \frac{F(n)}{n} \ll \exp\left(\sum_{p \le Z} \frac{F(p)}{p} - \frac{1}{10} R \log R\right),$$

uniformly for R satisfying $1 \le R \le \frac{\log Z}{\log\log Z}$. Here the thresshold for "sufficiently large", as well as the implied constant, depends at most on the constant A_1 and the function $A_2(\epsilon)$ associated to F in the definition of $\mathscr{M}$.

Let $F = fg$, $Z = x^{\theta}$, and $R = r$. Using that $f \in \mathscr{M}$, it is straightforward to check that $F \in \mathscr{M}$; moreover, the A_1 and $A_2(\epsilon)$ corresponding to f suffice to determine, together with k, choices for A_1 and $A_2(\epsilon)$ corresponding to F. By (2.4), $2 \le R \le \frac{\log Z}{\log\log Z}$. Applying Lemma 2.2,

$$\begin{aligned} \sum_{\substack{x^{\theta/2} < d \le x^{\theta} \\ P^+(d) \le x^{\theta/r}}} \frac{f(d)g(d)}{d} &\ll \exp\left(\sum_{p \le x^{\theta}} \frac{f(p)g(p)}{p} - \frac{1}{10} r \log r\right) \\ &\ll \exp\left(-\frac{1}{10} r \log r\right) \exp\left(\sum_{p \le x} \frac{f(p)}{p}\right). \end{aligned}$$

(We used once more than $f(p)g(p)/p = f(p)/p + O(1/p^2)$.) We put this estimate back into (2.5) and then sum on r in the range (2.4). Since

$$\sum_{r} \exp(O(r)) \exp(-\frac{1}{10} r \log r) \ll 1,$$

we conclude that the total contribution of n from category (3) is

$$\ll \left(\frac{y}{\log x} \prod_{p \le x} \left(1 - \frac{\nu(p)}{p}\right)\right) \exp\left(\sum_{p \le x} \frac{f(p)}{p}\right),$$

which is

$$\ll \frac{y}{\log x} \exp\left(\sum_{p \le x} \frac{f(p) - \nu(p)}{p}\right),$$

as desired. This completes the proof of Theorem 1.1.

3. Equidistribution of square roots of -1 modulo shifted primes: Proof of Theorem 1.3

We follow the original arguments of Hooley [Hoo64] as closely as possible.

For each positive integer k, let $\varrho(k)$ denote the number of square roots of -1 modulo k. From elementary number theory, ϱ is multiplicative and $\varrho(p^s) \le 2$ for all prime powers p^s, so that $\varrho(k) \le 2^{\omega(k)}$. Moreover, as noted in the introduction, $\varrho(k) = \frac{1}{4}r(k)$ for squarefree values of k. Thus, letting

$$\mathscr{K} = \{q-1 : q \text{ prime}\},$$

Theorem B′ gives that

$$\sum_{\substack{k\le x\\ k\in\mathscr{K}}} \varrho(k) \gg \frac{x}{\log x}. \tag{3.1}$$

For each pair of integers h, k with $k > 0$, let

$$S(h,k) = \sum_{\substack{\varpi \bmod k\\ \varpi^2\equiv -1 \pmod k}} e(h\varpi/k).$$

Trivially, $|S(h,k)| \le \varrho(k)$. By Weyl's criterion and the lower bound (3.1), to prove Theorem 1.3 it will suffice to show that

$$\sum_{\substack{k\le x\\ k\in\mathscr{K}}} S(h,k) = o(x/\log x) \qquad (x\to\infty)$$

for each fixed $h \ne 0$. (Cf. the discussion on p. 48 of [Hoo64].)

In what follows, we let

$$X = x^{1/\log\log x}.$$

For each positive integer $k \le x$, we write $k = k_1k_2$, where k_1 is X-smooth and k_2 is "X-rough" (meaning that $P^-(k_2) > X$). We decompose

$$\sum_{\substack{k\le x\\ k\in\mathscr{K}}} S(h,k) = \sum_{\substack{k\le x\\ k\in\mathscr{K}\\ k_1\le x^{1/3}}} S(h,k) + \sum_{\substack{k\le x\\ k\in\mathscr{K}\\ k_1> x^{1/3}}} S(h,k) = \sum\nolimits_1 + \sum\nolimits_2,$$

say. Concerning $\sum_2$, Cauchy–Schwarz gives

$$\Big|\sum\nolimits_2\Big| \le \Bigg(\sum_{\substack{k\le x\\ k_1> x^{1/3}}} 1\Bigg)^{1/2}\Bigg(\sum_{\substack{k\le x\\ k_1> x^{1/3}}} |S(h,k)|^2\Bigg)^{1/2}. \tag{3.2}$$

Since $|S(h,k)| \le \varrho(k) \le 2^{\omega(k)}$, we have that

$$\sum_{\substack{k\le x\\ k_1> x^{1/3}}} |S(h,k)|^2 \le \sum_{k\le x} 2^{2\omega(k)} \ll x(\log x)^3;$$

the final estimate here follows, for instance, from Shiu's Theorem A (or the much more elementary Theorem 01 in [HT88]). On the other hand, by the estimate for '$\Theta(x,y,z)$'

appearing at the bottom of p. 9 of [HT88],

$$\sum_{\substack{k\le x\\ k_1>x^{1/3}}} 1 \le x\exp\left(-\left(\frac{1}{3}+o(1)\right)\log\log x\cdot\log\log\log x\right),$$

as $x\to\infty$. (A less precise estimate, still sufficient here, follows from [BV64].) Putting these estimates back into (3.2), we see that

$$\sum\nolimits_2 = O(x/(\log x)^A)$$

for each fixed A. Thus, it will suffice to show that $\sum_1 = o(x/\log x)$, as $x\to\infty$.

As in [Hoo64] (see that paper's Lemma 3), we have $S(h,k) = S(h\bar{k}_2,k_1)S(h\bar{k}_1,k_2)$, where $\bar{k}_1$ denotes the inverse of k_1 modulo k_2, and $\bar{k}_2$ denotes the inverse of k_2 modulo k_1. Now using k_1 and k_2 for generic X-smooth and X-rough numbers,

$$\begin{aligned}\sum\nolimits_1 &= \sum_{\substack{k_1k_2\le x\\ k_1k_2\in\mathscr{K},\ k_1\le x^{1/3}}} S(h\bar{k}_2,k_1)S(h\bar{k}_1,k_2)\\ &\ll \sum_{\substack{k_1k_2\le x\\ k_1k_2\in\mathscr{K},\ k_1\le x^{1/3}}} \varrho(k_2)|S(h\bar{k}_2,k_1)| = \sum_{k_1\le x^{1/3}}\Theta(x/k_1,k_1),\end{aligned}$$

where, for $y\ge x^{2/3}$ and $k_1\le x^{1/3}$, we set

$$\Theta(y,k_1) = \sum_{\substack{k_2\le y\\ k_1k_2\in\mathscr{K}}} \varrho(k_2)|S(h\bar{k}_2,k_1)|.$$

Note that by Cauchy–Schwarz,

$$\Theta(y,k_1)^2 \le \Bigg(\sum_{\substack{k_2\le y\\ k_1k_2+1\text{ prime}}}\varrho(k_2)^2\Bigg)\Bigg(\sum_{\substack{k_2\le y\\ k_1k_2+1\text{ prime}}}|S(h\bar{k}_2,k_1)|^2\Bigg). \tag{3.3}$$

The first parenthesized sum in (3.3) can be handled by Theorem 1.1. With P the product of the primes not exceeding X, we have that

$$\sum_{\substack{k_2\le y\\ k_1k_2+1\text{ prime}}}\varrho(k_2)^2 \le \sum_{\substack{n\le y\\ k_1n+1\text{ prime}}}\mathbb{1}_{\gcd(n,P)=1}2^{2\omega(n)}.$$

To proceed, we observe that for k_1n+1 to be prime, either $n\le y^{1/2}$, or n avoids the class of $-k_1^{-1}$ mod p for all primes $p\le X$ not dividing k_1. The former case accounts for $O(y^{1/2})$ values of n. By Theorem 1.1, the latter case contributes

$$\ll \frac{y}{\log y}\exp\Bigg(\sum_{\substack{p\le X\\ p\nmid k_1}}\frac{-1}{p}+\sum_{X<p\le y}\frac{4}{p}\Bigg) \ll \frac{y}{\log y}\frac{1}{\log X}\left(\frac{\log y}{\log X}\right)^4\exp\Bigg(\sum_{p\mid k_1}\frac{1}{p}\Bigg) \ll \frac{y(\log\log x)^6}{(\log x)^2}$$

such values. (We used here that $\log y\asymp\log x$, that $\log X = \log x/\log\log x$, and that $\exp(\sum_{p\mid k_1}\frac{1}{p}) \ll \frac{k_1}{\varphi(k_1)} \ll \log\log 3k_1 \ll \log\log x$.) The contribution of $y^{1/2}$ is negligible compared to this, and so

$$\sum_{\substack{k_2 \le y \\ k_1k_2+1 \text{ prime}}} \varrho(k_2)^2 \ll \frac{y(\log\log x)^6}{(\log x)^2}.$$

Turning to the second parenthesized sum in (3.3), we have that

$$\sum_{\substack{k_2 \le y \\ k_1k_2+1 \text{ prime}}} |S(h\bar{k}_2, k_1)|^2 = \sum_{0 \le a < k_1} |S(ah, k_1)|^2 \sum_{\substack{k_2 \le y \\ k_2 \equiv \bar{a} \pmod{k_1} \\ k_1k_2+1 \text{ prime}}} 1.$$

Writing a_0 for the least nonnegative residue of $\bar{a}$ modulo k_1, we see that the values of k_2 appearing in the second sum have the form $a_0 + k_1 t$ for some $t \le y/k_1$. Moreover, either $t \le (y/k_1)^{1/2}$, or both both $a_0 + k_1 t$ and $k_1(a_0 + k_1 t) + 1$ have no prime factors exceeding X. Brun's sieve now implies that the sum on k_2 is

$$\ll \frac{y}{k_1(\log X)^2} \exp\left(\sum_{p \mid k_1} \frac{2}{p}\right) \ll \frac{y(\log\log x)^4}{k_1(\log x)^2},$$

and thus

$$\sum_{0 \le a < k_1} |S(ah, k_1)|^2 \sum_{\substack{k_2 \le y \\ k_2 \equiv \bar{a} \pmod{k_1} \\ k_1k_2+1 \text{ prime}}} 1 \ll \frac{y(\log\log x)^4}{k_1(\log x)^2} \sum_{0 \le a < k_1} |S(ah, k_1)|^2.$$

Exactly as in [Hoo64] (see Lemma 1 there), the sum on a is at most $\varrho(k_1)k_1 \gcd(h, k_1)$, and now collecting estimates yields

$$\sum_{\substack{k_2 \le y \\ k_1k_2+1 \text{ prime}}} |S(h\bar{k}_2, k_1)|^2 \ll \frac{y(\log\log x)^4}{(\log x)^2} \varrho(k_1)$$

where here and for the remainder of the argument, the implied constants may depend on h.

Referring back to (3.3),

$$\Theta(y, k_1) \ll \frac{y}{(\log x)^2} (\log\log x)^5 \varrho(k_1)^{1/2},$$

so that

$$\sum\nolimits_1 \ll \sum_{k_1 \le x^{1/3}} \Theta(x/k_1, k_1) \ll \frac{x}{(\log x)^2} (\log\log x)^5 \sum_{k_1 \le x^{1/3}} \frac{\varrho(k_1)^{1/2}}{k_1}.$$

Bounding the sum on k_1 by an Euler product, as in [Hoo64] (cf. the display immediately preceding that paper's eq. (12)), we find that

$$\sum_{k_1 \le x^{1/3}} \frac{\varrho(k_1)^{1/2}}{k_1} \ll (\log x)^{1/\sqrt{2}}.$$

We conclude that

$$\sum\nolimits_1 \ll \frac{x}{(\log x)^{2 - \frac{1}{\sqrt{2}}}} (\log\log x)^5.$$

Since $2 - \frac{1}{\sqrt{2}} > 1$, this implies that $\sum_1 = o(x/\log x)$, as desired. This completes the proof of Theorem 1.3.

Acknowledgements

The author is supported by NSF award DMS-1402268. He would like to express his gratitude to Terence Tao for a helpful blog post [Tao11] explaining Erdős's proof in [Erd52].

References

[Bar64a] M. B. Barban, *Multiplicative functions of* $^{\Sigma}R-$*equidistributed sequences*, Izv. Akad. Nauk UzSSR Ser. Fiz.–Mat. Nauk **1964** (1964), no. 6, 13–19 (Russian).

[Bar64b] ———, *On the number of divisors of "translations" of the prime number-twins*, Acta Math. Acad. Sci. Hungar. **15** (1964), 285–288 (Russian).

[BV64] M. B. Barban and A. I. Vinogradov, *On the number-theoretic basis of probabilistic number theory*, Dokl. Akad. Nauk SSSR **154** (1964), 495–496 (Russian).

[BV69] M. B. Barban and P. P. Vehov, *Summation of multiplicative functions of polynomials*, Mat. Zametki **5** (1969), 669–680 (Russian).

[DFI95] W. Duke, J. B. Friedlander, and H. Iwaniec, *Equidistribution of roots of a quadratic congruence to prime moduli*, Ann. of Math. (2) **141** (1995), 423–441.

[EH66] P. D. T. A. Elliott and H. Halberstam, *Some applications of Bombieri's theorem*, Mathematika **13** (1966), 196–203.

[Erd52] P. Erdős, *On the sum* $\sum_{k=1}^{x} d(f(k))$, J. London Math. Soc. **27** (1952), 7–15.

[Hoo57] C. Hooley, *On the representation of a number as the sum of two squares and a prime*, Acta Math. **97** (1957), 189–210.

[Hoo64] ———, *On the distribution of the roots of polynomial congruences*, Mathematika **11** (1964), 39–49.

[HT88] R. R. Hall and G. Tenenbaum, *Divisors*, Cambridge Tracts in Mathematics, vol. 90, Cambridge University Press, Cambridge, 1988.

[Kát68] I. Kátai, *A note on a sieve method*, Publ. Math. Debrecen **15** (1968), 69–73.

[Kát70] ———, *On an application of the large sieve: Shifted prime numbers, which have no prime divisors from a given arithmetical progression*, Acta Math. Acad. Sci. Hungar **21** (1970), 151–173.

[Lin60] Yu. V. Linnik, *An asymptotic formula in the Hardy–Littlewood additive problem*, Izv. Akad. Nauk SSSR, Ser. Mat. **24** (1960), 629–706 (Russian).

[MV07] H. L. Montgomery and R. C. Vaughan, *Multiplicative number theory. I. Classical theory*, Cambridge Studies in Advanced Mathematics, vol. 97, Cambridge University Press, Cambridge, 2007.

[NT98] M. Nair and G. Tenenbaum, *Short sums of certain arithmetic functions*, Acta Math. **180** (1998), 119–144.

[Shi80] P. Shiu, *A Brun-Titchmarsh theorem for multiplicative functions*, J. Reine Angew. Math. **313** (1980), 161–170.

[Tao11] T. Tao, *Erdős's divisor bound*, 2011, blog post published July 23, 2011 at `http://terrytao.wordpress.com/2011/07/23/erdos-divisor-bound/`. To appear in the forthcoming book *Spending symmetry*.

[VL57] A. I. Vinogradov and Yu. V. Linnik, *Estimate of the sum of the number of divisors in a short segment of an arithmetic progression*, Uspehi Mat. Nauk (N.S.) **12** (1957), no. 4(76), 277–280 (Russian).

[Wol71] D. Wolke, *Multiplikative Funktionen auf schnell wachsenden Folgen*, J. Reine Angew. Math. **251** (1971), 54–67.

Department of Mathematics, University of Georgia, Athens, GA 30602
E-mail address: `pollack@uga.edu`