Proc. Indian Acad. Sci. (Math. Sci.) Vol. 122, No. 1, February 2012, pp. 1–13.

On a paper of S S Pillai

M RAM MURTY[1] and R THANGADURAI[2]

[1]Department of Mathematics, Queen's University, Kingston, Ontario K7L 3N6, Canada
[2]Harish-Chandra Research Institute, Chhatnag Road, Jhusi, Allahabad 211 019, India
E-mail: murty@mast.queensu.ca; thanga@hri.res.in

MS received 9 January 2011; revised 20 April 2011

Abstract. In 1935, Erdös proved that all natural numbers can be written as a sum of a square of a prime and a square-free number. In 1939, Pillai derived an asymptotic formula for the number of such representations. The mathematical review of Pillai's paper stated that the proof of the above result contained inaccuracies, thus casting a doubt on the correctness of the paper. In this paper, we re-examine Pillai's paper and show that his argument was essentially correct. Afterwards, we improve the error term in Pillai's theorem using the Bombieri–Vinogradov theorem.

Keywords. Additive problems; applications of Bombieri–Vinogradov theorem.

1. Introduction

Let n be an natural number. Define

$$R(n) := \#\{n = p^2 + f : p \text{ is a prime and } f \text{ is a square-free integer}\}.$$

Thus $R(n)$ is the number of representations of n as a sum of a square of a prime and a square-free integer. Erdös [2] proved that $R(n) > 0$, when $n \not\equiv 1 \pmod 4$. It is easy to see that $R(n) = 0$ if $n \equiv 1 \pmod 4$ since every odd square is congruent to 1 (mod 4) which forces f to be divisible by 4. In 1939, Pillai [6] (see page 253 of [1]) provided an asymptotic formula for $R(n)$ as follows:

If $n \not\equiv 1 \pmod 4$,

$$R(n) \sim \frac{2\sqrt{n}}{\log n} \prod_q \left(1 - \frac{2}{q(q-1)}\right),$$

where the product runs through all the primes for which n is a quadratic residue.

Scherk [8], in his short review of Pillai's paper wrote: "Let $R(n)$ denote the number of representations of n as the sum of the square of a prime and a square-free integer; $n \not\equiv 1 \pmod 4$. Following Erdös's proof of $R(n) > 0$ [*J. London Math. Soc.* **10** (1935) 243–245], the author proves

$$R(n) = \frac{2\sqrt{n}}{\log n} \prod_q \left(1 - \frac{2}{q(q-1)}\right) + O\left(\frac{\sqrt{n}}{(\log n)(\log\log n)}\right),$$

where q runs through all the primes for which n is a quadratic residue. The paper contains inaccuracies." The review does not indicate where the inaccuracies are, or if the inaccuracies are major or minor. The reader would get the impression that Pillai's paper was flawed. A more accurate assessment of his paper would be that it was poorly written with inadequate references.

In this paper, we show Pillai's argument is essentially correct. Also, his proof contains the essence of the 'simple asymptotic sieve' later developed by Hooley [5], in his conditional proof of the Artin's primitive root conjecture. Moreover, by applying the Bombieri–Vinogradov theorem, we will improve the error term of Pillai. More precisely, we prove:

Theorem 1 [6]. *For any integer $n \not\equiv 1 \pmod 4$, we have*

$$R(n) = \mathrm{li}(\sqrt{n}) \prod_{\substack{q \\ \left(\frac{n}{q}\right)=1}} \left(1 - \frac{2}{q(q-1)}\right) + O\left(\frac{\sqrt{n}}{(\log n)(\log\log n)}\right),$$

where $\mathrm{li}(x) = \displaystyle\int_2^x \frac{\mathrm{d}t}{\log t}$.

Let us note that for any natural number N,

$$\mathrm{li}(x) = \frac{x}{\log x} \sum_{k=0}^{N} \frac{k!}{\log^k x} + O\left(\frac{x}{(\log x)^{N+1}}\right),$$

a fact easily verified by integrating by parts.

Theorem 2. *For any integer $n \not\equiv 1 \pmod 4$ and for any given $A \geq 1$, we have*

$$R(n) = \mathrm{li}(\sqrt{n}) \prod_{\substack{q \\ \left(\frac{n}{q}\right)=1}} \left(1 - \frac{2}{q(q-1)}\right) + O\left(\frac{\sqrt{n}}{(\log n)^A}\right).$$

In the concluding remarks, we consider the more general problem of representing natural number as the sum of the k-th power of a prime and a k-free integer. We indicate how the techniques of this paper can be used to treat this problem.

2. Preliminaries

Throughout the paper, we will use the fundamental relation

$$\sum_{d^2 | n} \mu(d) = \begin{cases} 1, & \text{if } n \text{ is square free,} \\ 0, & \text{otherwise.} \end{cases}$$

For given natural numbers a and d such that $(a, d) = 1$, we denote by $\pi(x, d, a)$ the number of primes $p \equiv a \pmod d$ such that $p \leq x$. We define

$$E(x, d) := \max_{(a,d)=1} \left| \pi(x, d, a) - \frac{\mathrm{li}(x)}{\phi(d)} \right|.$$

We also denote by $\nu(d)$, the number of distinct prime factors of an integer $d \geq 2$. It is well-known that

$$\nu(d) \ll \frac{\log n}{\log \log n}.$$

Theorem 3 (The Siegel–Walfisz theorem). *For any $A > 0$ and $B > 0$, we have that*

$$\pi(x, d, a) = \frac{\mathrm{li}(x)}{\phi(d)} + O\left(\frac{x}{(\log x)^A}\right)$$

holds uniformly for all $d \leq (\log x)^B$.

Theorem 4 (The Bombieri–Vinogradov theorem). *For any $A > 0$ there exists $B = B(A) > 0$ such that*

$$\sum_{d \leq \frac{\sqrt{x}}{(\log x)^B}} E(x, d) \ll \frac{x}{(\log x)^A}.$$

For our applications, we need the following weighted version of the Bombieri–Vinogradov theorem.

Theorem 5. *Let $C \geq 1$ be a real number. For any $A > 0$, there exists $B = B(A, C)$ such that*

$$\sum_{d \leq \frac{\sqrt{x}}{(\log x)^B}} C^{\nu(d)} E(x, d) \ll \frac{x}{(\log x)^A}.$$

Proof. Apply Theorem 4 with $2A$ to get a constant B. Note that, as $\pi(x, d, a) \leq 2x/d$ for all $d \leq x$, we have

$$E(x, d) \ll \frac{x}{d}.$$

Therefore, by putting $z = \dfrac{\sqrt{x}}{(\log x)^B}$, we have

$$\begin{aligned} \sum_{d \leq z} C^{\nu(d)} E(x, d) &\ll \sum_{d \leq z} C^{\nu(d)} \left(\frac{x}{d}\right)^{1/2} (E(x, d))^{1/2} \\ &\ll \sqrt{x} \left(\sum_{d \leq z} \frac{C^{2\nu(d)}}{d}\right)^{1/2} \left(\sum_{d \leq z} E(x, d)\right)^{1/2}, \end{aligned}$$

by the Cauchy–Schwarz inequality. By Theorem 4 and by the choice $2A$, we see that

$$\left(\sum_{d \leq z} E(x, d)\right)^{1/2} \ll \frac{\sqrt{x}}{(\log x)^A}.$$

Considering the other sum, we have by standard methods that (see, for example [7])

$$\left(\sum_{d\leq z}\frac{C^{2\nu(d)}}{d}\right)^{1/2} \ll (\log z)^{C^2} \ll (\log x)^{C^2}.$$

Thus, we get the required estimate. □

In our discussion below, we will use the following standard estimates:

$$\sum_{p>y}\frac{1}{p(p-1)} \ll \frac{1}{y}; \quad \sum_{d|P_y} 2^{\nu(d)} \ll 3^y, \quad \text{where } P_y = \prod_{p\leq y} p.$$

As in the papers of Erdös [2] and (implicit in) Pillai [6], we need the following result.

Theorem 6 [3]. *Let $A > 0$ and $B > 0$ be integers. Then the number of integer solutions for $n = Ax^2 + By^2$ is $\leq 2d(n)$, where $d(n)$ is the number of divisors of n.*

Since we indicate in the last section that Pillai's method extends to count the number of representations of a natural number as a sum of a k-th power of a prime and a k-free integer, we record here the neccesary generalization of Theorem 2 required in the proof. This result can be found in the paper by Evelyn and Linfoot [4] where it is written that the argument originates with H. Rademacher and was extended by A. Oppenheim.

Theorem 7. *Let a, b be positive integers. The number of solutions of*

$$ax^k + by^k = n$$

is bounded by $(k(k-1)+1)d_{k^2}(n)$ where $d_t(n)$ is the number of ways of writing n as a product of t natural numbers.

For other standard estimates in analytic number theory, we refer to [7].

3. Proof of Theorem 1

Let us first note that

$$R(n) = S_1 + S_2, \tag{1}$$

where

$$S_1 = \#\{n = p^2 + f : p \text{ is a prime and } (p, n) = 1,\ f \text{ is a square-free integer}\} \tag{2}$$

and

$$S_2 = \#\{n = p^2 + f : p \text{ is a prime and } p|n,\ f \text{ is a square-free integer}\}. \tag{3}$$

Since $S_2 \leq \nu(n)$, the number of distinct prime factors of n and $\nu(n) = O\left(\frac{\log n}{\log\log n}\right)$, we conclude that

$$S_2 = O\left(\frac{\log n}{\log\log n}\right). \tag{4}$$

Hence to compute $R(n)$, we need to essentially compute S_1.

It is clear from Pillai's paper that one can apply the simple asymptotic sieve to treat S_1. Indeed, let $N(n, y)$ (with y to be chosen later) be the number of primes $p \leq \sqrt{n}$ such that $n - p^2$ is not divisible by a square of a prime q with $q < y$.

Thus, we see that

$$S_1 \leq \#\{p \leq \sqrt{n} : q \text{ prime such that } q^2 | (n - p^2) \Longrightarrow q > y\} := N(n, y). \tag{5}$$

On the other hand, we,

$$S_1 \geq N(n, y) - \#\{p \leq \sqrt{n} : \exists q > y \text{ prime such that } q^2 | (n - p^2)\}. \tag{6}$$

From (5) and (6), we will show that $N(n, y)$ gives the main term and the rest is an error term.

To treat $N(n, y)$, let $P(k)$ denote the largest prime factor of k and $\mu(k)$ the Mobius function. Then clearly,

$$\begin{aligned} N(n, y) &= \sum_{\substack{p \leq \sqrt{n} \\ (p,n)=1}} \sum_{\substack{d^2 | (n-p^2) \\ P(d) \leq y}} \mu(d) \\ &= \sum_{\substack{d \leq \sqrt{n} \\ P(d) \leq y \\ (d,n)=1}} \mu(d) \sum_{\substack{p \leq \sqrt{n} \\ p^2 \equiv n \pmod{d^2}}} 1. \end{aligned}$$

For a given natural number n, let $\rho(d^2) = \#\left\{1 \leq x \leq d^2 : x^2 \equiv n \pmod{d^2}\right\}$. By the Chinese Remainder theorem, $\rho(d^2)$ is a multiplicative function of d and $\rho(d^2) = O(2^{\nu(d)})$. Then as $(n, d) = 1$, we have for each such x by Theorem 3, for any $A, B > 0$,

$$\frac{\mathrm{li}(\sqrt{n})}{\phi(d^2)} + O\left(\frac{\sqrt{n}}{(\log n)^A}\right) \tag{7}$$

number of primes $p \leq n$ satisfying $p^2 \equiv n \pmod{d^2}$ uniformly for all $d \leq (\log n)^B$. Therefore, we get

$$\begin{aligned} N(n, y) &= \sum_{\substack{d \leq \sqrt{n} \\ P(d) \leq y \\ (d,n)=1}} \mu(d) \left[\frac{\rho(d^2)\mathrm{li}(\sqrt{n})}{\phi(d^2)} + O\left(\frac{\rho(d^2)\sqrt{n}}{(\log n)^A}\right)\right]. \\ &= \mathrm{li}(\sqrt{n}) \sum_{\substack{d \leq \sqrt{n} \\ P(d) \leq y \\ (d,n)=1}} \frac{\mu(d)\rho(d^2)}{\phi(d^2)} + O\left(\sum_{\substack{d \leq \sqrt{n} \\ P(d) \leq y \\ (d,n)=1}} \frac{|\mu(d)|\rho(d^2)\sqrt{n}}{(\log n)^A}\right). \\ &= T_1 + T_2 \quad \text{(say)}. \end{aligned}$$

Note that

$$T_1 = \mathrm{li}(\sqrt{n}) \sum_{\substack{d=1 \\ P(d) \leq y \\ (d,n)=1}}^{\infty} \frac{\mu(d)\rho(d^2)}{\phi(d^2)} - \mathrm{li}(\sqrt{n}) \sum_{\substack{d > \sqrt{n} \\ P(d) \leq y \\ (d,n)=1}}^{\infty} \frac{\mu(d)\rho(d^2)}{\phi(d^2)}.$$

Let us write

$$A(n) = \prod_{p,(p,n)=1} \left(1 - \frac{\rho(p^2)}{p(p-1)}\right).$$

By the multiplicativity of $\rho(d^2)$, we have

$$\begin{aligned}
\sum_{\substack{d=1 \\ P(d)\le y \\ (d,n)=1}}^{\infty} \frac{\mu(d)\rho(d^2)}{\phi(d^2)} &= \prod_{\substack{p\le y \\ (p,n)=1}} \left(1 - \frac{\rho(p^2)}{p(p-1)}\right) \\
&= \prod_{\substack{p \\ (p,n)=1}} \left(1 - \frac{\rho(p^2)}{p(p-1)}\right) \prod_{\substack{p>y \\ (p,n)=1}} \left(1 - \frac{\rho(p^2)}{p(p-1)}\right)^{-1} \\
&= A(n) \prod_{\substack{p>y \\ \left(\frac{n}{p}\right)=1}} \left(1 + \frac{2/(p(p-1)}{1 - \frac{2}{p(p-1)}}\right) \\
&= A(n) \prod_{\substack{p>y \\ \left(\frac{n}{p}\right)=1}} \left(1 + O\left(\frac{1}{p(p-1)}\right)\right) \\
&= A(n) \prod_{\substack{p>y \\ \left(\frac{n}{p}\right)=1}} \exp\left(\frac{O(1)}{p(p-1)}\right) \\
&= A(n) \exp\left(\sum_{p>y} O\left(\frac{1}{p(p-1)}\right)\right) \\
&= A(n) \exp\left(O\left(\frac{1}{y}\right)\right) = A(n)\left(1 + O\left(\frac{1}{y}\right)\right).
\end{aligned}$$

Now,

$$\begin{aligned}
-\mathrm{li}(\sqrt{n}) \sum_{\substack{d>\sqrt{n} \\ P(d)\le y \\ (d,n)=1}}^{\infty} \frac{\mu(d)\rho(d^2)}{\phi(d^2)} &= O\left(\mathrm{li}(\sqrt{\mathrm{n}}) \sum_{d>\sqrt{n}} \frac{2^{\nu(d)}}{d^2}\right) \\
&= O\left(\mathrm{li}(\sqrt{\mathrm{n}}) \int_{\sqrt{n}}^{\infty} \frac{S(m)}{m^3} dm\right), \\
&\qquad \text{where } S(m) = \sum_{d\le m} 2^{\nu(d)} = O(m \log m) \\
&= O\left(\mathrm{li}(\sqrt{n}) \frac{\log n}{\sqrt{n}}\right) = O(1).
\end{aligned}$$

Therefore,

$$T_1 = \mathrm{li}(\sqrt{n}) \prod_{\substack{p \\ \left(\frac{n}{p}\right)=1}} \left(1 - \frac{2}{p(p-1)}\right)\left(1 + O\left(\frac{1}{y}\right)\right) + O\,(1)\,.$$

Consider

$$\begin{aligned} T_2 &= O\left(\frac{\sqrt{n}}{(\log n)^A} \sum_{\substack{d \le \sqrt{n} \\ P(d) \le y}} |\mu(d)| 2^{\nu(d)}\right) \\ &= O\left(\frac{\sqrt{n}}{(\log n)^A} \sum_{d | P_y} 2^{\nu(d)}\right), \quad \text{where } P_y = \prod_{p \le y} p \\ &= O\left(\frac{3^y \sqrt{n}}{(\log n)^A}\right) \end{aligned}$$

Thus,

$$\begin{aligned} N(n, y) &= \mathrm{li}(\sqrt{n}) \prod_{\substack{p \\ \left(\frac{n}{p}\right)=1}} \left(1 - \frac{2}{p(p-1)}\right)\left(1 + O\left(\frac{1}{y}\right)\right) \\ &\quad + O\,(1) + O\left(\frac{3^y \sqrt{n}}{(\log n)^A}\right). \end{aligned} \tag{8}$$

Consider now

$$M(y, n) = \#\left\{p \le \sqrt{n} : \exists q > y \text{ prime such that } q^2 | (n - p^2)\right\}.$$

Following Pillai's outline, we compute $M(y, n)$ in three intervals, namely:

(1) $y < q < (\log n)^c$;
(2) $(\log n)^c < q < \frac{\sqrt{n}}{(\log n)^c}$; and
(3) $\frac{\sqrt{n}}{(\log n)^c} < q < \sqrt{n}$, where c is a suitable constant to be chosen later.

By Theorem 3, we have

$$\begin{aligned} M_1(y, n) &= \sum_{\substack{y < q < (\log n)^c \\ (q,n)=1}} \sum_{\substack{p \le \sqrt{n} \\ p^2 \equiv n \pmod{q^2} \\ (p,q)=1}} 1 \\ &= \sum_{\substack{y < q < (\log n)^c \\ (q,n)=1}} \left[\frac{\rho(q^2)\mathrm{li}(\sqrt{n})}{q(q-1)} + O\left(\frac{\rho(q^2)\sqrt{n}}{(\log n)^A}\right)\right] \\ &\le \mathrm{li}(\sqrt{n}) \sum_{y < q} \frac{2}{q(q-1)} + O\left(\frac{\sqrt{n}}{(\log n)^A} \sum_{y < q < (\log n)^c} \rho(q^2)\right). \end{aligned}$$

Therefore,

$$M_1(y,n) = O\left(\frac{\sqrt{n}}{(\log n)} \times \frac{1}{y}\right) + O\left(\frac{\sqrt{n}\log\log n}{(\log n)^{A-c}}\right). \tag{9}$$

Now, consider

$$\begin{aligned}
M_2(y,n) &= \sum_{\substack{(\log n)^c < q < \frac{\sqrt{n}}{(\log n)^c} \\ (q,n)=1}} \sum_{\substack{p \le \sqrt{n} \\ p^2 \equiv n \pmod{q^2} \\ (p,q)=1}} 1 \\
&\le \sum_{\substack{(\log n)^c < q < \frac{\sqrt{n}}{(\log n)^c} \\ (q,n)=1}} \sum_{\substack{p \le \sqrt{n} \\ \frac{n-p^2}{q^2} = k}} 1 \\
&\le \sum_{\substack{(\log n)^c < q < \frac{\sqrt{n}}{(\log n)^c} \\ (q,n)=1}} \sum_{\substack{a \le \sqrt{n} \\ \frac{n-a^2}{q^2} = k}} 1 \\
&= \sum_{\substack{(\log n)^c < q < \frac{\sqrt{n}}{(\log n)^c} \\ (q,n)=1}} \left(\frac{2\sqrt{n}}{q^2} + 1\right) \\
&\le 2\sqrt{n} \sum_{q > (\log n)^c} \frac{1}{q(q-1)} + \frac{\sqrt{n}}{(\log n)^c} \\
&\le \frac{3\sqrt{n}}{(\log n)^c}.
\end{aligned}$$

Therefore

$$M_2(y,n) = O\left(\frac{\sqrt{n}}{(\log n)^c}\right). \tag{10}$$

Finally, consider

$$\begin{aligned}
M_3(y,n) &= \sum_{\substack{\frac{\sqrt{n}}{(\log n)^c} < q < \sqrt{n} \\ (q,n)=1}} \sum_{\substack{p \le \sqrt{n} \\ p^2 \equiv n \pmod{q^2}}} 1 \\
&= \sum_{\substack{\frac{\sqrt{n}}{(\log n)^c} < q < \sqrt{n} \\ (q,n)=1}} \sum_{\substack{p \le \sqrt{n} \\ n = p^2 + kq^2}} 1.
\end{aligned}$$

Thus, $M_3(y,n)$ counts the number of triples (k, p, q) satisfying the conditions stipulated in the summations. Since $q > \sqrt{n}/(\log n)^c$, we see that $q^2 > n/(\log n)^{2c}$. Since $kq^2 = n - p^2 < n$, we see that the integer k is at most $(\log n)^{2c}$. For each such integer k, we can have at most $2d(n)$ number of solutions of $n = p^2 + kq^2$, by Theorem 6. Therefore, we have

$$M_3(y,n) = O((\log n)^{2c} d(n)) = O((\log n)^{2c} n^{\epsilon}) \qquad \text{for any } \epsilon > 0. \tag{11}$$

Thus, from (9), (10) and (11), we get

$$M(y,n) = O\left(\frac{\sqrt{n}}{(\log n)y}\right) + O\left(\frac{\sqrt{n}}{(\log n)^{A-c}}\right) + O\left(\frac{\sqrt{n}}{(\log n)^{c}}\right) + O\left((\log n)^{2c} n^{\epsilon}\right). \tag{12}$$

Choose $y = c_1 \log\log n$, $A = 4$ and $c = 2$, where $c_1 > 0$ is a fixed constant. With these choices, we see that

$$M(y,n) = O\left(\frac{\sqrt{n}}{(\log n)(\log\log n)}\right).$$

Therefore, by (8) we get

$$N(n,y) = \mathrm{li}(\sqrt{n}) \prod_{\substack{p \\ \left(\frac{n}{p}\right)=1}} \left(1 - \frac{2}{p(p-1)}\right) + O\left(\frac{\sqrt{n}}{(\log n)(\log\log n)}\right).$$

Thus, we arrive at

$$R(n) = \mathrm{li}(\sqrt{n}) \prod_{\substack{p \\ \left(\frac{n}{p}\right)=1}} \left(1 - \frac{2}{p(p-1)}\right) + O\left(\frac{\sqrt{n}}{(\log n)(\log\log n)}\right).$$

This proves the theorem. □

4. Proof of Theorem 2

We shall not apply the asymptotic sieve. Instead, we compute the formula directly.

$$\begin{aligned} R(n) &= \sum_{p^2 \le n} \sum_{d^2 \mid (n-p^2)} \mu(d) \\ &= \sum_{p^2 \le n} \left(\sum_{\substack{d^2 \mid (n-p^2) \\ d<z}} \mu(d) + \sum_{\substack{d^2 \mid (n-p^2) \\ d \ge z}} \mu(d) \right) \end{aligned}$$

for a suitable z to be chosen later. For a given $A \ge 1$, we choose $B > 0$ as in Theorem 2.3. Consider

$$\sum_{d<z} \mu(d) \sum_{\substack{p \le \sqrt{n} \\ p^2 \equiv n \pmod{d^2}}} 1 = \sum_{d<z} \mu(d) F(\sqrt{n}, d^2),$$

where

$$F(\sqrt{n}, d^2) = \begin{cases} \dfrac{\rho(d^2)\mathrm{li}(\sqrt{n})}{\phi(d^2)} + O\left(\rho(d^2)E(\sqrt{n}, d^2)\right), & \text{if } d < n^{1/4}/(\log n)^B, \\ \leq \rho(d^2)\left(\dfrac{\sqrt{n}}{d^2} + 1\right), & \text{if } n^{1/4}/(\log n)^B < d < z. \end{cases}$$

Using the familiar estimate

$$\sum_{d<z} 2^{\nu(d)} \ll z \log z,$$

which implies by partial summation the estimate

$$\sum_{d>y} \frac{2^{\nu(d)}}{d^2} \ll \frac{\log y}{y},$$

we have

$$\begin{aligned} &\sum_{n^{1/4}/(\log n)^B < d < z} \mu(d)\rho(d^2)\left(\frac{\sqrt{n}}{d^2} + 1\right) \\ &\leq \sqrt{n} \sum_{n^{1/4}/(\log n)^B < d < z} \frac{\rho(d^2)}{d^2} + O(z \log z) \\ &\leq \sqrt{n} \sum_{n^{1/4}/(\log n)^B < d < z} \frac{2^{\nu(d)}}{d^2} + O(z \log z) \\ &= O(n^{1/4} \log^{B+1} n) + O(z \log z). \end{aligned}$$

We choose $z = \sqrt{n}/\log^{2A} n$ so that the error above is

$$O\left(\frac{\sqrt{n}}{\log^{2A-1} n}\right).$$

By Theorem 5, we have

$$\begin{aligned} &O\left(\sum_{d<n^{1/4}/(\log n)^B} |\mu(d)|\rho(d^2)E(\sqrt{n}, d^2)\right) \\ &= O\left(\sum_{d<n^{1/4}/(\log n)^B} |\mu(d)|\rho(d^2)E(\sqrt{n}, d^2)\right) \\ &= O\left(\frac{\sqrt{n}}{(\log n)^A}\right). \end{aligned}$$

Therefore, since $A \geq 1$,

$$\begin{aligned}\sum_{d<z} \mu(d) F(\sqrt{n}, d^2) &= \sum_{d<n^{1/4}/(\log n)^B} \frac{\mu(d)\rho(d^2)\,\mathrm{li}(\sqrt{n})}{\phi(d^2)} \\ &\quad + O\left(\frac{\sqrt{n}}{(\log n)^A}\right) \\ &= \mathrm{li}(\sqrt{n}) \sum_{d=1}^{\infty} \frac{\mu(d)\rho(d^2)}{\phi(d^2)} \\ &\quad - \mathrm{li}(\sqrt{n}) \sum_{d>n^{1/4}/(\log n)^B} \frac{\mu(d)\rho(d^2)}{\phi(d^2)} \\ &\quad + O\left(\frac{\sqrt{n}}{(\log n)^A}\right).\end{aligned}$$

Using the familiar inequality $\phi(d) \gg d/\log\log d$, we see from partial integration that

$$\sum_{d>y} \frac{2^{\nu(d)} \log\log d}{d^2} \ll \frac{\log y \log\log y}{y}$$

so that the second sum above is bounded by

$$O\left(\frac{\log^{B+1} n}{n^{1/4}}\right).$$

Thus,

$$\begin{aligned}\sum_{d<z} \mu(d) F(\sqrt{n}, d^2) &= \mathrm{li}(\sqrt{n}) \sum_{d=1}^{\infty} \frac{\mu(d)\rho(d^2)}{\phi(d^2)} + O\left(\frac{\sqrt{n}}{(\log n)^A}\right) \\ &= \mathrm{li}(\sqrt{n}) \prod_{\substack{q \\ \left(\frac{n}{q}\right)=1}} \left(1 - \frac{2}{q(q-1)}\right) + O\left(\frac{\sqrt{n}}{(\log n)^A}\right).\end{aligned}$$

This proves that

$$\sum_{d<z} \mu(d) F(\sqrt{n}, d^2) = \mathrm{li}(\sqrt{n}) \prod_{\substack{q \\ \left(\frac{n}{q}\right)=1}} \left(1 - \frac{2}{q(q-1)}\right) + O\left(\frac{\sqrt{n}}{(\log n)^A}\right).$$

Now, consider

$$\left| \sum_{p\leq\sqrt{n}} \sum_{\substack{d>z \\ n=p^2+kd^2}} \mu(d) \right| \leq \sum_{k<n/z^2} \#\{(d,p) : d^2k + p^2 = n\}.$$

By Theorem 6, we get

$$\sum_{p\leq\sqrt{n}} \sum_{\substack{d>z\\ n=p^2+kd^2}} \mu(d) \leq \sum_{k\leq n/z^2} 2d(k) \ll n^{\epsilon} n/z^2 \leq n^{\epsilon}(\log n)^A,$$

where $0 < \epsilon < 1/2$. Thus,

$$R(n) = \mathrm{li}(\sqrt{n}) \prod_{\substack{q\\ \left(\frac{n}{q}\right)=1}} \left(1 - \frac{2}{q(q-1)}\right) + O\left(\frac{\sqrt{n}}{(\log n)^A}\right).$$

This proves our theorem. □

5. Concluding remarks

Let $k \geq 2$ be an integer. Let

$$R(n,k) := \#\{n = p^k + f : p \text{ is a prime, } f \text{ is a k-free integer}\}.$$

Then, Erdös proved that $R(n,k) > 0$. One may get a similar asymptotic formula as for $R(n,k)$ by imitating our proof of Theorem 2. Hence we omit the proof of the following theorem.

Theorem 8. *Let $\rho(p^k)$ denote the number of solutions of the congruence $x^k \equiv n \pmod{p^k}$. We have for any $A > 0$,*

$$R(n,k) = \mathrm{li}(n^{1/k}) \prod_{\substack{q\\ \left(\frac{n}{q}\right)=1}} \left(1 - \frac{\rho(p^k)}{q^{k-1}(q-1)}\right) + O\left(\frac{n^{1/k}}{(\log n)^A}\right).$$

The question of whether these error terms can be improved is an interesting one. Without going into details, we remark that one can use Montgomery's conjecture on the error term

$$E(n,d) = O((n/d)^{1/2} d^{\epsilon}),$$

for any $\epsilon > 0$ to deduce that the error terms in Theorem 2 can be improved to $O(n^{\theta})$ for some $\theta < 1/2$. We leave the details to the reader.

Acknowledgements

We thank the referee for helpful comments on an earlier version of this paper. We also thank Prof. R Balasubramanian for some useful discussions in the writing of this paper. The first author is grateful to the Institute for Mathematical Sciences (Chennai) for the salubrious environment in which this collaboration took place. The research of the first author was partially supported by an NSERC Discovery Grant.

References

[1] Balasubramanian R and Thangadurai R (eds), Collected works of S. Sivasankaranarayana Pillai, Ramanujan Mathematical Society (2010) No. 1, pp. 253–255

[2] Erdös P, The representation of an integer as the sum of the square of a prime and of a square-free integer, *J. London Math. Soc.* **10** (1935) 243–245

[3] Estermann T, Einige Satze uber quadratfreie Zahlen, *Math. Ann.* **105** (1931) 653–662

[4] Evelyn C J A and Linfoot E H, On a problem in the additive theory of numbers (second paper), *J. Reine Angew. Math.* **164** (1931) 131–140

[5] Hooley C, Artin's conjecture for primitive roots, *J. Reine Angew. Math.* **225** (1967) 209–220

[6] Pillai S S, On the number of representation of a number as the sum of the square of a prime and a square-free integers, *Proc. Indian Acad. Sci. (Math. Sci.)* **10** (1939) 390–391

[7] Ram Murty M, Problems in Analytic Number Theory, Graduate Texts in Mathematics, 2nd edition (Springer) (2008)

[8] Scherk P, Mathematical Reviews MR0000839 (1135h)