Riv. Mat. Univ. Parma (7) **3*** (2004), 205-212

GIOVANNI COPPOLA (*)

# On the symmetry of square-free numbers in almost all short intervals (**)

## 1 - Introduction and statement of the results

In this paper we study the symmetry, in almost all short intervals, of the square-free numbers.

As in the previous papers [**2**], [**3**], [**4**] and [**5**] on the symmetry of (respectively) the function $\omega(n)$ (i.e. the prime-divisors function), $\Lambda(n)$ (i.e. the von Mangoldt function), $d(n)$ (the number of divisors of $n$) and a wide class of arithmetical functions, we apply the Large Sieve for, say,

$$\chi_q(x) \stackrel{def}{=} \left\{ \frac{x+h}{q} \right\} + \left\{ \frac{x-h}{q} \right\} - 2\left\{ \frac{x}{q} \right\};$$

this periodic function checks the symmetry of our arithmetical function $f$ ($f = \mu^2$ in the present paper) in almost all short intervals.

As usual, we say that something holds in «almost all short intervals» $[x-h, x+h]$, as $N \leq x \leq 2N$, if it is true in the «short» intervals $[x-h, x+h]$, i.e. $h = h(N) \to \infty$ and $h = o(N)$ (as $N \to \infty$), with a number of «exceptions» $x \in [N, 2N]$ which is $o(N)$, as $N \to \infty$.

In this and the quoted papers we write $f$ as a Dirichlet convolution, involving the constant function **1**; whence the modules $q$ appear (in the following $q$ and $d$ will be called «divisors»). Then the Large Sieve is applied.

---

(*) DIIMA, Università di Salerno, Via Ponte Don Melillo, 84084 Fisciano (SA) - Italy; e-mail: gcoppola@diima.unisa.it

(**) Received February 23rd 2004. AMS classification 11 N 25, 11 N 36, 11 N 37.

However, a novelty with respect to our previous papers (in particular, w.r.t. [**4**]) is due to the very low density of these divisors for $\mu^2$; in fact, we easily obtain a first non-trivial upper bound (which is uniform w.r.t. $h$, see Theorem 1); and also, using the Large Sieve (like in [**4**]), an essentially optimal upper bound (see Theorem 2 and Corollary 2).

In fact, the key point is the low density of the divisors, not that of the original sequence (which is well-known to be positive and easy to calculate). It turns out that this property is used in Lemma 2, which gives an essentially optimal treatment of the «sporadic» terms.

We explicitly remark that our arguments are elementary, using only the Large Sieve.

We plan to continue our study by means of deeper methods.
First of all, let's define the «symmetry-sum» for square-free numbers

$$S^{\pm}(x) \overset{def}{=} \sum_{|n-x| \leq h} \operatorname{sgn}(n-x)\,\mu^2(n)\,.$$

The mean-square over $x \sim N$ (i.e., $N < x \leq 2N$) will be consequently:

$$I(N, h) \overset{def}{=} \sum_{x \sim N} |S^{\pm}(x)|^2.$$

We need, then, $\chi_q(x)$ finite Fourier expansion (here, of course, $c_{q,q} = 0$):

$$c_{j,q} \overset{def}{=} = \frac{1}{q} \sum_{r \leq q} \chi_q(r)\, e_q(-rj) \Rightarrow \chi_q(x) = \sum_{j<q} c_{j,q}\, e_q(jx) = \sum_{d|q} \frac{d}{q} \sum_{j \leq d}{}^{*} c_{j,d}\, e_d(jx)$$

(here and in the sequel the $*$ means, as usual, a sum over reduced classes); this is proved using the property $c_{dj',dq'} = \dfrac{1}{d} c_{j',q'}$, $\forall d, j', q' \in \boldsymbol{N}$ (since $qc_{j,q}$ depends only on $j/q$); the coefficients $c_{j,q}$ satisfy ($\| \ \|$ is the distance from $\boldsymbol{Z}$):

$$\sum_{j \leq q} |c_{j,q}|^2 = \frac{2}{q} \sum_{r \leq |h| (\operatorname{mod} q)} \chi_q^2(x) = 2 \left\| \frac{h}{q} \right\|.$$

All these properties of $\chi_q$ will be useful in the sequel (in Lemma 3) to apply the Large Sieve (Lemma 1).

We remark that all our variables will be integer (actually, natural numbers).

In the sequel we will write $L \overset{def}{=} \log N$.

We start giving a first upper bound for $I(N, h)$, by our

Theorem 1. *Let* $N, h$ *be natural numbers, with* $h = h(N) \to \infty$ *and* $h = o(N)$ *as* $N \to \infty$. *Then*

$$\sum_{x \sim N} \left| \sum_{|n-x| \leq h} \mu^2(n) \operatorname{sgn}(n-x) \right|^2 \ll NhL^2.$$

We write «a.a.» for «almost all», i.e. all, but $o(N)$ possible exceptions.
As an immediate consequence of Theorem 1, we then get our

Corollary 1. *Let* $N, h$ *be natural numbers, with* $h = h(N) \to \infty$ *and* $h = o(N)$ *as* $N \to \infty$. *Then for any positive increasing function* $\eta(N) \to \infty$ *as* $N \to \infty$

$$\left| \sum_{|n-x| \leq h} \mu^2(n) \operatorname{sgn}(n-x) \right| \ll \eta(N) \sqrt{h} \log x, \qquad \text{a.a. } x \in [N, 2N].$$

We now come to more precise results, achieved by a careful use of the Large Sieve, together with an essentially optimal bound for the «sporadic» terms; this is provided by our

Theorem 2. *Let* $N, h$ *be natural numbers, with* $h = h(N) \to \infty$ *as* $N \to \infty$ *and* $h \leq N^{\vartheta}$, $0 < \vartheta < 1/2$. *Then*

$$I(N, h) \ll N\sqrt{h} \left( 1 + \left( \frac{h^{15}}{N^4} \right)^{1/14} N^{\varepsilon} \right).$$

As an easy consequence, we get our

Corollary 2. *Let* $N, h \in \mathbf{N}$, *with* $h = h(N) \to \infty$ *as* $N \to \infty$ *and* $h \leq N^{4/15 - \varepsilon}$. *Then*

$$I(N, h) \ll N\sqrt{h},$$

*whence for any positive increasing function* $\eta(N) \to \infty$ *as* $N \to \infty$

$$|S^{\pm}(x)| \ll \eta(N) h^{1/4}, \qquad \text{a.a. } x \in [N, 2N].$$

As regards the «Selberg integral» for $\mu^2$, i.e.

$$J(N, h) \overset{def}{=} \sum_{x \sim N} \left| \sum_{x < n \leq x+h} \mu^2(n) - \frac{6}{\pi^2} h \right|^2$$

it is trivially true that $I(N, h) \ll J(N, h) + h^3$.

Hall [6] has estimated $J$ (hence, also $I$) in the range $h = o(N^{2/9} L^{-4/9})$.

We explicitly remark that our estimates supersede Hall's, in the full range; also, ours hold in the wider range $h \leqslant N^{4/15 - \varepsilon}$.

Furthermore, our results are proved by elementary means (i.e. the Large Sieve), while his arguments involve complex integration and Dirichlet polynomials estimates.

The paper is organized as follows:

– in section 2 we give the necessary Lemmas;

– in section 3 we apply them to prove our Theorems and Corollaries.

The Author wishes to thank Professor Saverio Salerno and Professor Alberto Perelli for friendly and helpful comments.

## 2 - Lemmas

Lemma 1. *Let $Q$ and $N$ be natural numbers, $M$ be an integer and $\lambda_{a,q}$ be complex numbers* ($\forall a, q \in \mathbf{N}$); *then*

$$\sum_{n=M+1}^{M+N} \left| \sum_{q \leqslant 2Q} \sum_{a \leqslant q}{}^{*} \lambda_{a,q} e_q(an) \right|^2 \ll (N + Q^2) \sum_{q \leqslant 2Q} \sum_{a \leqslant q}{}^{*} |\lambda_{a,q}|^2 .$$

This is obtained by duality from the well-known Large Sieve inequality [1].

Lemma 2. *Let $N$, $h$ and $D$ be natural numbers, with $h = h(N) \to \infty$ and $h = o(N)$ as $N \to \infty$; suppose that $D = \infty(\sqrt{h})$ and $D \ll \sqrt{N}$; then*

$$\sum_{x \sim N} \left| \sum_{d \sim D} \mu(d)\, \chi_{d^2}(x) \right|^2 \ll \frac{N h^2 N^{\varepsilon}}{D} .$$

Proof. We use an argument similar to the proof of Theorem 1:

$$\sum_{x \sim N} \left| \sum_{d \sim D} \mu(d)\, \chi_{d^2}(x) \right|^2 \ll \sum_{d_1, d_2 \sim D} \sum_{x \sim N} \sum_{\frac{x-h}{d_1^2} \leqslant m_1 \leqslant \frac{x+h}{d_1^2}} \sum_{\frac{x-h}{d_2^2} \leqslant m_2 \leqslant \frac{x+h}{d_2^2}} 1$$

$$\ll \sum_{d_1, d_2 \sim D} \sum_{\frac{N-h}{d_1^2} \leqslant m_1 \leqslant \frac{2N+h}{d_1^2}} \sum_{\substack{\frac{N-h}{d_2^2} \leqslant m_2 \leqslant \frac{2N+h}{d_2^2} \\ \frac{m_1 d_1^2 - 2h}{d_2^2} \leqslant m_2 \leqslant \frac{m_1 d_1^2 + 2h}{d_2^2}}} h$$

$$\ll h \sum_{d_1 \sim D} \sum_{\frac{N-h}{d_1^2} \leqslant m_1 \leqslant \frac{2N+h}{d_1^2}} \sum_{|n - m_1 d_1^2| \leqslant 2h} d(n) ,$$

where $d(n)$ is the (usual) divisor function; hence:

$$\sum_{x\sim N}\left|\sum_{d\sim D}\mu(d)\,\chi_{d^2}(x)\right|^2 \ll Nh^2\,\frac{N^{\varepsilon}}{D},$$

by the classical estimate $d(n)\ll n^{\varepsilon}$. ■

Lemma 3. *Let $N, h\in\mathbf{N}$, with $h=h(N)\to\infty$ as $N\to\infty$ and $h\leqslant N^{\vartheta}$, $0<\vartheta<1/2$. Then*

$$\sum_{x\sim N}\left|\sum_{q\leqslant J}\mu(q)\,\chi_{q^2}(x)\right|^2 \ll (N+J^{5/2}L^2)\sqrt{h},$$

*provided $J=\infty(\sqrt{h})$ and $J=o(\sqrt{N})$.*

Proof. Indicating (as in Theorem 2 proof) $\Sigma_1(x)=\sum_{q\leqslant J}\mu(q)\,\chi_{q^2}(x)$, and writing (as before)

$$\chi_{q^2}(x)=\sum_{d|q^2}\frac{d}{q^2}\sum_{j<d}{}^{*}\,c_{j,\,d}\,e_d(jx),$$

we get, setting $q=smk$ and $d=sm^2$, with $\mu^2(s)=\mu^2(m)=\mu^2(k)=1$, $(s,m)=1$ and $(k,sm^2)=1$:

$$\Sigma_1(x)=\sum_{d\leqslant J^2}\left(\sum_{\substack{q\leqslant J\\ q^2\equiv 0(\mathrm{mod}\,d)}}\mu(q)\,\frac{d}{q^2}\right)\sum_{j<d}{}^{*}\,c_{j,\,d}\,e_d(jx)$$

$$=\sum_{s\leqslant J}\frac{\mu(s)}{s}\sum_{\substack{m\leqslant J/\sqrt{s}\\ (m,s)=1}}\mu(m)\sum_{\substack{k\leqslant J\\ (k,sm^2)=1}}\frac{\mu(k)}{k^2}\sum_{j<sm^2}{}^{*}\,c_{j,\,sm^2}\,e_{sm^2}(jx).$$

Hence, by a dyadic dissection argument, we can confine to:

$$\Sigma_1(x,S)\stackrel{def}{=}\sum_{s\sim S}\frac{\mu(s)}{s}\sum_{\substack{m\leqslant J/\sqrt{s}\\ (m,s)=1}}\mu(m)\,\alpha_{sm^2}(J)\sum_{j<sm^2}{}^{*}\,c_{j,\,sm^2}\,e_{sm^2}(jx),$$

where, say:

$$\alpha_n(J)\stackrel{def}{=}\sum_{\substack{k\leqslant J\\ (k,n)=1}}\frac{\mu(k)}{k^2}\ll\sum_{k\leqslant J}\frac{1}{k^2}\ll 1\,;$$

then, by the Large Sieve (Lemma 1)

$$\sum_{x\sim N}|\Sigma_1(x,S)|^2 \ll (N+S^4)\sum_{s\sim S}\frac{1}{s^2}\sum_{m\leq\frac{J}{\sqrt{s}}}\left\|\frac{h}{sm^2}\right\|$$

$$\ll (N+S^4)\left(\sum_{s\sim S}\frac{1}{s^2}\sqrt{\frac{h}{s}}+h\sum_{s\sim S}\frac{1}{s^3}\sum_{\sqrt{\frac{2h}{s}}<m\leq\frac{J}{\sqrt{s}}}\frac{1}{m^2}\right)\ll\frac{N+S^4}{S^{3/2}}\sqrt{h}\,,$$

having split as $sm^2\leq 2h$ and $sm^2>2h$ for the distance to the integers. Since the $S\ll N^\varepsilon$ contribute a negligible amount to $\Sigma_1(x)$ mean-square,

$$\Sigma_1(x)\ll L\max_{N^\varepsilon\ll S\ll J}|\Sigma_1(x,S)|\Rightarrow\sum_{x\sim N}|\Sigma_1(x)|^2\ll L^2\max_{N^\varepsilon\ll S\ll J}\frac{N+S^4}{S^{3/2}}\sqrt{h}\,,$$

whence the Lemma. ■

## 3 - Proof of the theorems and of the corollaries

We start proving Theorem 1 (and Corollary 1 follows immediately from it). First of all, $\mu^2(n)=\sum_{d^2|n}\mu(d)$ (see [7]), whence $S^{\pm}(x)$ equals

$$\sum_{|n-x|\leq h}\mu^2(n)\,\mathrm{sgn}\,(n-x)=-\sum_{d\leq\sqrt{x+h}}\mu(d)\,\chi_{d^2}(x)+\mathcal{O}\Big(\sum_{d^2|x}1+\sum_{d^2|x-h}1\Big).$$

Hence (ignoring, in the sequel, these remainders, since their mean-square contributes $\mathcal{O}(N)$, a negligible amount):

$$\sum_{|n-x|\leq h}\mu^2(n)\,\mathrm{sgn}\,(n-x)=-\sum_{d\leq\sqrt{2N+h}}\mu(d)\,\chi_{d^2}(x)\,,$$

because $q>x+h\Rightarrow\chi_q(x)=0$ (see its definition). We get

$$S^{\pm}(x)=-\sum_{d\leq\sqrt{2h}}\mu(d)\,\chi_{d^2}(x)-\sum_{\sqrt{2h}<d\leq\sqrt{2N+h}}\mu(d)\,\chi_{d^2}(x)\,,$$

where the first sum contains **non-sporadic** terms (i.e. $\left[\frac{x-h}{d^2},\frac{x+h}{d^2}\right]$ contains more than one integer) and the second consists of **sporadic** terms (i.e. $\left[\frac{x-h}{d^2},\frac{x+h}{d^2}\right]$ contains at most one integer). Non-sporadic terms are estimated trivially by $\chi_q(x) \ll 1$ (fractional parts are $\mathcal{O}(1)$)

$$\sum_{d \leqslant \sqrt{2h}} \mu(d)\, \chi_{d^2}(x) \ll \sqrt{h}$$

**individually**, i.e. $\forall x \in [N, 2N]$. (Its mean-square contributes $\mathcal{O}(Nh)$.) We estimate trivially the sporadic terms mean-square (after a dyadic dissection):

$$\sum_{x \sim N} \Big| \sum_{\sqrt{2h} < d \leqslant \sqrt{2N+h}} \mu(d)\, \chi_{d^2}(x) \Big|^2 \ll L^2 \max_{\sqrt{h} \ll D \ll \sqrt{N}} \sum_{x \sim N} \Big| \sum_{d \sim D} \mu(d)\, \chi_{d^2}(x) \Big|^2$$

$$\ll L^2 \max_{\sqrt{h} \ll D \ll \sqrt{N}} Nh \ll NhL^2,$$

since $\chi_{d^2}(x) \ll \sum_{\frac{x-h}{d^2} \leqslant m \leqslant \frac{x+h}{d^2}} 1$ gives (in the stated range we have $\frac{h}{d^2} \ll 1$)

$$\sum_{x \sim N} \Big| \sum_{d \sim D} \mu(d)\, \chi_{d^2}(x) \Big|^2 \ll \sum_{d_1, d_2 \sim D} \sum_{x \sim N} \sum_{\frac{x-h}{d_1^2} \leqslant m_1 \leqslant \frac{x+h}{d_1^2}} \sum_{\frac{x-h}{d_2^2} \leqslant m_2 \leqslant \frac{x+h}{d_2^2}} 1$$

$$\ll \sum_{d_1, d_2 \sim D} \sum_{\frac{N-h}{d_1^2} \leqslant m_1 \leqslant \frac{2N+h}{d_1^2}} \sum_{\substack{\frac{N-h}{d_2^2} \leqslant m_2 \leqslant \frac{2N+h}{d_2^2} \\ \frac{m_1 d_1^2 - 2h}{d_2^2} \leqslant m_2 \leqslant \frac{m_1 d_1^2 + 2h}{d_2^2}}} h \ll h \sum_{d_1 \sim D} \sum_{d_2 \sim D} \frac{N}{d_2^2} \ll Nh. \quad \blacksquare$$

We now prove Theorem 2.

We follow the first steps of the previous proof and split at $J = J(N, h)$:

$$S^{\pm}(x) = -\sum_{d \leqslant J} \mu(d)\, \chi_{d^2}(x) - \sum_{J < d \leqslant \sqrt{2N+h}} \mu(d)\, \chi_{d^2}(x) = -\Sigma_1(x) - \Sigma_2(x),$$

say, where $\Sigma_1(x) \overset{def}{=} \sum_{d \leqslant J} \mu(d)\, \chi_{d^2}(x)$, $\Sigma_2(x) \overset{def}{=} \sum_{J < d \leqslant \sqrt{2N+h}} \mu(d)\, \chi_{d^2}(x)$.

Let's treat first the «high divisors», by Lemma 2:

$$\sum_{x \sim N} |\Sigma_2(x)|^2 \ll L^2 \max_{J \ll D \ll \sqrt{N}} \sum_{x \sim N} \Big| \sum_{d \sim D} \mu(d)\, \chi_{d^2}(x) \Big|^2 \ll \frac{Nh^2 N^{\varepsilon}}{J}.$$

Hence, we only need Lemma 3 for the remaining «low divisors». In fact,

$$\sum_{x \sim N} |\Sigma_1(x)|^2 \ll (N + J^{5/2} L^2) \sqrt{h}.$$

Gathering the estimates for the mean-squares of $\Sigma_1(x)$ and $\Sigma_2(x)$, together with the optimal choice $J = N^{2/7} h^{3/7} N^{\varepsilon}$ (with $\varepsilon > 0$ small, which respects the hypotheses of Lemmas 2 and 3), we get the Theorem. $\blacksquare$

We then get immediately Corollary 2. $\blacksquare$

## References

[1] E. BOMBIERI, *Le grand crible dans la théorie analytique des nombres*, Astérisque 18, Societè Mathematique de France, Paris 1974. MR 51 #8057

[2] G. COPPOLA, *On the symmetry of distribution of the prime-divisors function in almost all short intervals*, to appear.

[3] G. COPPOLA, *On the symmetry of primes in almost all short intervals*, Ricerche Mat. **52** (2003), 21-29.

[4] G. COPPOLA and S. SALERNO, *On the symmetry of the divisor function in almost all short intervals*, Acta Arith. **113** (2004), 189-201.

[5] G. COPPOLA and S. SALERNO, *On the symmetry of arithmetical functions in almost all short intervals*, C.R. Math. Acad. Sci. to appear.

[6] R. R. HALL, *Squarefree numbers on short intervals*, Mathematika **29** (1982), 7-17. MR 83m:10077

[7] G. TENENBAUM, *Introduction to analytic and probabilistic number theory*, Cambridge Studies in Advanced Mathematics, 46, Cambridge University Press, Cambridge 1995. MR 97e:11005b

## Abstract

*In this paper we study the symmetry of the square-free numbers in almost all short intervals; by elementary methods (based on the Large Sieve) we give an upper bound for the mean-square (over the interval $N < x \leq 2N$) of the «symmetry sum» $\sum_{|n-x| \leq h} sgn(n-x)\mu^2(n)$ ($sgn(t) := t/|t|$, $sgn(0) = 0$).*

* * *