

ON THE AVERAGE AND NORMAL ORDERS OF HOOLEY'S Δ -FUNCTION

R. R. HALL AND G. TENENBAUM

Introduction

In a recent paper [7], Hooley defined the functions

$$\Delta(n) = \max_u \text{card} \{d : d|n, u < d \leq eu\} \quad \text{and} \quad S(x) = \sum_{n < x} \Delta(n).$$

Hooley proved that $S(x) \ll x(\log x)^\lambda$ with $\lambda = (4/\pi) - 1 = 0.27323\dots$ and gave a proof, due to Erdős, that $S(x)/x \rightarrow \infty$ as $x \rightarrow \infty$. There followed applications of the former result to a problem in Diophantine approximation and to Waring's problem. In each of these applications (Theorems 4 and 5), and also in Theorem 3, an immediate improvement would be entailed by a reduction in the value of λ above. There were also applications to some divisor problems. Hooley expressed the belief that the true order of magnitude of $S(x)$ is roughly $x \log^A x$ where A is a positive absolute constant not substantially exceeded by $(4/\pi) - 1$.

In this paper we study both the average and normal orders of $\Delta(n)$. Regarding the former, we are able to improve on the results of Hooley and Erdős.

THEOREM 1. *We have $S(x) \ll x(\log x)^{A_0} \sqrt{\log \log x}$ where*

$$A_0 = \frac{1}{8} \{ (9 - \sqrt{(19/3)})(1 + \sqrt{(19/27)})^{1/3} + (9 + \sqrt{(19/3)})(1 - \sqrt{(19/27)})^{1/3} - 37/3 \} \\ \leq 0.23457.$$

THEOREM 2. *Let A be the infimum of the numbers λ for which $S(x) \ll_\lambda x(\log x)^\lambda$. Then $A < A_0$; indeed $A < 0.23454$.*

The improvement obtained in Theorem 2 is slight indeed, but the proof involves new ideas which may be significant elsewhere. In particular we are obliged to study the expected distribution of the prime factors of integers belonging to certain rather thin sequences. We make a conjecture later in the paper which would imply that $A \leq \sqrt{3} - 3/2$.

THEOREM 3. *We have $S(x) \gg x \log \log x$.*

This is better than the quantitative result we could obtain from Erdős' method, but it surely falls well short of the truth: it seems likely that

$$S(x) \sim xL(x) \log^A x$$

where $A > 0$ and L is a slowly oscillating function of $\log x$, that is $L(x) \sim L(x^t)$ for each fixed $t > 0$.

Received 9 January, 1981.

[J. LONDON MATH. SOC. (2), 25 (1982), 392-406]

We now turn to the normal order of $\Delta(n)$. We have the following.

THEOREM 4. *Let B be the infimum of the numbers λ such that for almost all n , we have $\Delta(n) < (\log n)^\lambda$. Then*

$$B \leq \log 2 - \frac{\log 2}{\log 3}.$$

We prove our results in order of difficulty.

Proof of Theorem 3. Let $T(n, \alpha)$ be the function defined in [4], that is

$$T(n, \alpha) = \text{card} \{d, d' : d|n, d'|n, |\log d/d'| \leq (\log n)^\alpha\}.$$

Then

$$T(n, 0) \leq \sum_{d|n} \text{card} \{d' : d/e \leq d' \leq de\} \leq 2\tau(n)\Delta(n).$$

It is known [6] that

$$\sum_{n < x} \frac{T(n, 0)}{2^{\Omega(n)}} \sim Cx \log \log x$$

where

$$C = \frac{2}{\pi} \prod_p \left(1 - \frac{1}{2p}\right)^{-4} \left(1 - \frac{1}{p}\right)^2 \left(1 - \frac{1}{4p^2}\right),$$

and since $\tau(n) \leq 2^{\Omega(n)}$ we have

$$S(x) \geq \frac{1}{2} \sum_{n < x} \frac{T(n, 0)}{\tau(n)} \geq (\frac{1}{2}C + o(1))x \log \log x.$$

This completes the proof.

Before embarking on the proof of Theorem 4 we establish the following lemmas, which are also needed for Theorem 2.

LEMMA 1. *Let V be a set of vertices of a k -dimensional hypercube and let $D(V)$ be the set of vectors \mathbf{d} of the form $\mathbf{d} = \mathbf{v} - \mathbf{v}'$, $\mathbf{v}, \mathbf{v}' \in V$. Then $|D(V)| \geq |V|^\beta$ where $\beta = (\log 3)/(\log 2)$.*

Proof. We apply the following theorem of Woodall [8]: if R, S are (not necessarily disjoint) sets of vertices and if $M = \{\mathbf{x} : \mathbf{x} = \mathbf{r} + \mathbf{s}, \mathbf{r} \in R, \mathbf{s} \in S\}$, then

$$|M| \geq (|R| |S|)^{\beta/2}.$$

Now put $R = V, S = -V = \{\mathbf{e} - \mathbf{v} : \mathbf{v} \in V\}$ where $\mathbf{e} = (1, 1, \dots, 1)$. Then $M = D(V) + \mathbf{e}$ and the result follows.

LEMMA 2. *Let $U(n, \alpha)$ be the function defined in [6], that is*

$$U(n, \alpha) = \text{card} \{d, d' : d|n, d'|n, (d, d') = 1, |\log d/d'| \leq (\log n)^\alpha\}.$$

Then for squarefree n , we have $(\Delta(n))^\beta \leq 3U(n, 0)$.

Proof. Let $\Delta_1(n)$ be defined similarly to $\Delta(n)$, but with \sqrt{e} taking the place of e . Plainly $\Delta(n) \leq 2\Delta_1(n)$ so that it will be sufficient to prove that $(\Delta_1(n))^\beta \leq U(n, 0)$. Let $\Delta_1(n) = r$, so that for some u , n has r divisors such that $u < d_1 < d_2 < \dots < d_r < u\sqrt{e}$. Let $n = p_1 p_2 \dots p_k$ and with each divisor d associate the vector $\mathbf{v} = \{\varepsilon_1, \varepsilon_2, \dots, \varepsilon_k\}$ where $d = p_1^{\varepsilon_1} p_2^{\varepsilon_2} \dots p_k^{\varepsilon_k}$. These vectors are vertices of the k -dimensional hypercube and we suppose that V is the set of vertices associated with d_1, d_2, \dots, d_r so that $D(V)$ corresponds to ratios of these divisors. We deduce from Lemma 1 that there are at least r^β distinct ratios $d_i/d_j = \delta_{ij}/\delta'_{ij}$ say, where $(\delta_{ij}, \delta'_{ij}) = 1$ and $|\log \delta_{ij}/\delta'_{ij}| \leq 1$. So $U(n, 0) \geq r^\beta$. This completes the proof.

We remark that the exponent β is best possible. For let $n = p_1 p_2 \dots p_{2h}$ where $p_1 < p_2 < \dots < p_{2h} < (1 + 1/h)p_1$. Every divisor with h prime factors lies in the interval $(p_1^h, p_1^h e)$ and so

$$\Delta(n) \geq \binom{2h}{h} >> \frac{2^{2h}}{\sqrt{h}}, \quad U(n, 0) \leq \sum_{l=0}^h \binom{2h}{2l} \binom{2l}{l} << \frac{3^{2h}}{\sqrt{h}}.$$

Therefore $U(n, 0) << (\Delta(n))^\beta (\log \Delta(n))^{(\beta-1)/2}$ for infinitely many n .

LEMMA 3. Let $\omega(n)$ and $\Omega(n)$ denote respectively the number of distinct prime factors of n and the total number of prime factors of n . Then for each fixed positive y we have that

$$\sum_{n < x} y^{\omega(n)} <<_y x (\log x)^{y-1};$$

moreover this holds with ω replaced by Ω on the left, provided that $y < 2$.

We omit the proof as asymptotic formulae for both sums are well known.

LEMMA 4. Let $Q(x, m) = \text{card} \left\{ n : n < x, \prod_{p|n} p = m \right\}$. Then $\sum_{m < x} (Q(x, m))^t <<_t x$ for each fixed positive t .

This is [5; Lemma 3]; the idea of the proof is due to Erdős.

Proof of Theorem 4. By [6; Théorème 3] we have that $U(m, 0) <<_\varepsilon (\log m)^{\log 3 - 1 + \varepsilon}$ for almost all m , and hence for almost all squarefree m . By Lemma 2 we also have that

$$\Delta(m) <<_\varepsilon (\log m)^{\log 2 - (\log 2/\log 3) + \varepsilon}$$

for almost all squarefree m . For any integer n and prime p we have $\Delta(np) \leq 2\Delta(n)$, and so for all n we have $\Delta(n) \leq 2^{\Omega(n) - \omega(n)} \Delta(m)$ where m is the product of the distinct prime factors of n . Set

$$\lambda = \log 2 - \frac{\log 2}{\log 3} + 3\varepsilon.$$

We are going to show that the number of integers $n < x$ such that $\Delta(n) \geq (\log n)^\lambda$ is

$o(x)$. For each such n , at least one of the inequalities

$$2^{\Omega(n)-\omega(n)} \geq (\log n)^\epsilon, \quad \Delta(m) \geq (\log m)^{\lambda-\epsilon}$$

holds. Since

$$\sum_{n < x} \{\Omega(n) - \omega(n)\} \ll x$$

the number of $n < x$ for which the first inequality holds is $o(x)$. The number of $n < x$ for which the second inequality holds is

$$\begin{aligned} &\leq \sum_{m < x} \{Q(x, m) : \Delta(m) \geq (\log m)^{\lambda-\epsilon}\} \\ &\leq \left(\sum_{m < x} Q^2(x, m) \right)^{1/2} \left(\sum_{m < x} \{1 : \Delta(m) \geq (\log m)^{\lambda-\epsilon}\} \right)^{1/2} \end{aligned}$$

by the Cauchy-Schwarz inequality. Since $\Delta(m) \ll_\epsilon (\log m)^{\lambda-2\epsilon}$ for almost all m , we also have

$$\sum_{m < x} \{1 : \Delta(m) \geq (\log m)^{\lambda-\epsilon}\} = o(x),$$

and, by Lemma 4 with $t = 2$, the sum in the first bracket above is $\ll x$. Hence the second inequality holds for only $o(x)$ integers $n < x$, and this completes the proof.

Next, we require some more information about the function $U(n, 0)$.

LEMMA 5. For each fixed $y \in (0, 2)$ we have that

$$\sum_{n < x} U(n, 0) y^{\Omega(n)} < <_y \begin{cases} x(\log x)^{y-1} & \text{if } 0 < y < \frac{1}{2}, \\ x(\log x)^{-1/2} \log \log x & \text{if } y = \frac{1}{2}, \\ x(\log x)^{3y-2} & \text{if } \frac{1}{2} < y < 2. \end{cases}$$

Proof. This is not quite contained in [6; Théorème 2]. But the left hand side does not exceed

$$\sum_{n < x} \sum'_{dd'|n} y^{\Omega(n)} \leq \sum'_{dd' < x} y^{\Omega(dd')} \sum_{m < x/dd'} y^{\Omega(m)}$$

where \sum' denotes summation restricted by the condition that $|\log d/d'| \leq 1$. This is

$$\begin{aligned} &< <_y \sum'_{dd' < x} y^{\Omega(dd')} \frac{x}{dd'} \left(\log \frac{x}{dd'} \right)^{y-1} < <_y \sum_{r < x} (\log r)^{y-1} \sum'_{dd' < x/r} y^{\Omega(dd')} \\ &< <_y \sum_{r < x} (\log r)^{y-1} \left(\sum_{d < 2\sqrt{(x/r)}} y^{\Omega(d)} \right)^2 < <_y \sum_{r < x} \frac{x}{r} (\log r)^{y-1} \left(\log \frac{x}{r} \right)^{2y-2}, \end{aligned}$$

where $\log \zeta$ is to be interpreted as $\max(\log \zeta, 1)$ here and throughout the paper. The result follows.

The same inequalities hold if Ω is replaced by ω ; in this form the upper range for y is $(\frac{1}{2}, \infty)$. The proof is similar but slightly more awkward and we omit it.

LEMMA 6. For fixed r , $1 \leq r < (\log 3)/(\log 3/2)$, we have that

$$\sum_{n < x} \{U(n, 0)\}^{1/r} \ll x(\log x)^{3^{1/r}-1-1/r}.$$

Remark (i). If we set $r = \beta = (\log 3)/(\log 2)$ we obtain

$$\sum_{m < x} |\mu(m)|\Delta(m) \ll x(\log x)^{1-1/\beta}.$$

This leads to a proof that $A \leq 1 - 1/\beta$, but this result is inferior to Hooley's.

Remark (ii). For each n , define $\beta(n) = \log U(n, 0)/\log \Delta(n)$. For squarefree n , we know that $\beta(n) \geq \beta$. We conjecture that on a suitable sequence of asymptotic density 1 we have $\beta(n) \rightarrow 2$, which would of course imply that $B \leq \frac{1}{2}(\log 3 - 1)$. It may be true that

$$S(x) \ll_{\epsilon} \sum_{n < x} \{U(n, 0)\}^{1/2+\epsilon}$$

and this would give $A \leq \sqrt{3} - 3/2$.

Proof. Set $y = (1/3)^{1-1/r} > \frac{1}{2}$. By Hölder's inequality, we have

$$\sum_{n < x} \{U(n, 0)\}^{1/r} \leq \left(\sum_{n < x} U(n, 0)y^{\omega(n)} \right)^{1/r} \left(\sum_{n < x} (1/y)^{\omega(n)/(r-1)} \right)^{1-1/r}.$$

By Lemma 5, we have

$$\sum_{n < x} U(n, 0)y^{\omega(n)} \ll_y x(\log x)^{3y-2} \ll_r x(\log x)^{3^{1/r}-2}.$$

Next, $(1/y)^{\omega(n)/(r-1)} = 3^{1/r}$ and so

$$\sum_{n < x} (1/y)^{\omega(n)/(r-1)} \ll_r x(\log x)^{3^{1/r}-1}.$$

This gives the result stated.

Proof of Theorem 1. Our proof is based on the inequality

$$\Delta^2(n) \leq \sum'_{d|n} \Delta\left(\frac{n}{dd'}\right), \tag{1}$$

where summation is over all pairs of divisors d, d' such that $dd'|n$ and, as before, the dash denotes the further condition that $|\log d/d'| \leq 1$. To see this, let $\Delta(n) = r$ so that for some u , n has r divisors such that $u < d_1 < d_2 < \dots < d_r \leq eu$. For each i, j , $1 \leq i \leq r$, $1 \leq j \leq r$, consider the divisors d_i, d_j and set $d = d_{ij} = d_i/(d_i, d_j)$ and $d' = d'_{ij} = d_j/(d_i, d_j)$. Then $(d, d') = 1$ so that $dd'|n$; moreover $|\log d/d'| \leq 1$. Given

d, d' satisfying these conditions, we have to consider with how many pairs d_i, d_j they are associated. For some t we have $d_i = dt, d_j = d't$ where $t|n/dd'$; moreover $u < dt \leq eu, u < d't \leq eu$. Either of these last conditions, say the one involving d , places t in a fixed range $v < t \leq ev$ (in this case $v = u/d$) and so the number of choices for t is at most $\Delta(n/dd')$. This establishes the above inequality. We observe that the right hand side is at least $U(n, 0)$ so that our upper estimate for $S(x)$ will be at least equal to $\sum_{n < x} \{U(n, 0)\}^{1/2}$. In this connection we remark that this does not

necessarily put a lower limit $\sqrt{3} - 3/2$ on our estimate for A since the use of Hölder's inequality in Lemma 6 may be inefficient.

Next, we set out to prove the most precise inequalities we can obtain of the form

$$S(x, y) = \sum_{n < x} \Delta(n)y^{\Omega(n)} <<_y x(\log x)^{f(y)}(\log \log x)^{g(y)}$$

for each fixed y in the range $0 < y < 2$; that is, we try to find the smallest function $f(y)$ for which the above inequality holds. We are primarily interested in estimating $f(1)$, but our method involves an iteration and we are obliged to consider other values of y . Since $\Delta(n) \leq \tau(n)$ the inequality is valid with $f(y) = 2y - 1, g(y) = 0$; we could improve the coefficient of y here to $4/\pi$ using Hooley's method. Instead, we begin by observing that a permissible f is the function $f_1(y) = y - 1 + \chi(2y - 1)$, where here and elsewhere $\chi(t)$ denotes $\max(0, t)$, and $g_1(y) = 1$ or 0 according as $y = 1/2$ or not. For $\Delta(n) \leq U(n, 0)$ for all n : we notice that if $u < d_1 < d_2 < \dots < d_r \leq eu$ as before then each pair of (relatively prime) divisors $d_{1j} = d_1/(d_1, d_j), d'_{1j} = d_j/(d_1, d_j)$ ($1 \leq j \leq r$) contributes to U . We define the subsidiary sum

$$T(x, z) = \sum_{n < x} \Delta^2(n)z^{\Omega(n)},$$

and we begin the iteration. By the fundamental inequality (1) we have

$$\begin{aligned} T(x, z) &\leq \sum_{n < x} z^{\Omega(n)} \sum'_{dd'|n} \Delta\left(\frac{n}{dd'}\right) \leq \sum'_{dd' < x} z^{\Omega(dd')} S\left(\frac{x}{dd'}, z\right) \\ &<<_z (\log \log x)^{g_1(z)} \sum'_{dd' < x} z^{\Omega(dd')} \frac{x}{dd'} \left(\log \frac{x}{dd'}\right)^{f_1(z)} \\ &<<_z (\log \log x)^{g_1(z)} \sum'_{dd' < x} z^{\Omega(dd')} \sum_{m < x/dd'} (\log m)^{f_1(z)} \\ &<<_z (\log \log x)^{g_1(z)} \sum_{m < x} (\log m)^{f_1(z)} \sum'_{dd' < x/m} z^{\Omega(dd')} \\ &<<_z (\log \log x)^{g_1(z)} \sum_{m < x} (\log m)^{f_1(z)} \frac{x}{m} \left(\log \frac{x}{m}\right)^{2z-2}, \end{aligned}$$

the inner sum being estimated from above as in the proof of Lemma 5. Since $f_1 > -1$, this gives

$$T(x, z) <<_z x(\log x)^{f_1(z) + \chi(2z-1)} (\log \log x)^{2g_1(z)}.$$

Now let $y^2/2 < x < 2$. By the Cauchy-Schwarz inequality we deduce that

$$(S(x, y))^2 \ll_{y,z} T(x, z) \sum_{n < x} (y^2/z)^{\Omega(n)} \ll_{y,z} x(\log x)^{f_1(z) + \chi(2z-1) + y^2/z - 1} (\log \log x)^{2g_1(z)}.$$

We may therefore set

$$f_2(y) = \frac{1}{2} \min_z \left\{ f_1(z) + \chi(2z-1) + \frac{y^2}{z} - 1 : y^2/2 < z < 2 \right\}$$

$$= \begin{cases} y-1, & 0 < y \leq \frac{1}{2}, \\ y^2-3/4, & \frac{1}{2} < y \leq \frac{1}{2}\sqrt{5}, \\ y\sqrt{5}-2, & \frac{1}{2}\sqrt{5} < y < 2. \end{cases}$$

Since the optimal z satisfies

$$z = \begin{cases} y, & y \leq \frac{1}{2}, \\ \frac{1}{2}, & \frac{1}{2} \leq y \leq \frac{1}{2}\sqrt{5}, \\ y/\sqrt{5}, & y \geq \frac{1}{2}\sqrt{5}, \end{cases}$$

we find that $g_2(z) = 1$ or 0 according as $\frac{1}{2} \leq y \leq \frac{1}{2}\sqrt{5}$ or not. More generally, we may define the sequence of functions $f_i(y)$ by the formula

$$f_{i+1}(y) = \frac{1}{2} \min_z \left\{ f_i(z) + \chi(2z-1) + \frac{y^2}{z} - 1 : y^2/2 < z < 2 \right\}$$

with corresponding $g_i(y) : g_{i+1}(y) = \frac{1}{2}\{g_i(z) + g_1(z)\}$, this z being the optimum z occurring in the formula for $f_{i+1}(y)$. Notice that for $y \leq \frac{1}{2}$ there is no improvement from the iteration; indeed in this range f_1 is already sharp as $\Delta \geq 1$. Next, we find that

$$f_3(y) = \begin{cases} f_2(y), & 0 < y \leq \sqrt{3}/2, \\ z_0^2(y)/2 + z_0(y) + y^2/2z_0(y) - 11/8, & \sqrt{3}/2 < y < 2, \end{cases}$$

where $z_0(y)$ is the (real, positive) root of the equation $2z^2(z+1) = y^2$. Also, $g_3(y) = g_2(y)$ for $y \leq \sqrt{3}/2$; for $y > \sqrt{3}/2$ we have $g_3(y) = 1/2$. We substitute $y = 1$ and find that $f_3(1) = A_0$. This completes the proof.

We remark that further iterations lead to no better result. It is not difficult to prove, more generally, that if the definition of $S(x, y)$ is extended to all positive y (which requires that we replace Ω by ω) then for each fixed y we have $f_{i+1}(y) = f_i(y)$ for $i \geq i_0(y)$.

We require some further lemmas in preparation for the proof of Theorem 2. The first of these enables us to work with squarefree numbers only.

LEMMA 7. *Let A_1 denote the infimum of the numbers λ for which*

$$S_1(x) = \sum_{n < x} |\mu(m)| \Delta(m) \ll_{\lambda} x(\log x)^{\lambda}.$$

Then $A_1 = A$.

Proof. We just have to show that $A \leq A_1$. First, we observe that for any integer n and prime p (dividing n or not) we have $\Delta(p^n) \leq (t+1)\Delta(n)$. To see this, split the divisors δ of p^n into $t+1$ classes corresponding to the least integer s such that $\delta = p^s d$ where $d|n$. Then if $\delta|p^n$, $u < \delta \leq eu$, we have $d|n$, $up^{-s} < d \leq eup^{-s}$. This gives $t+1$ ranges for d ; in each range n has at most $\Delta(n)$ divisors. More generally, for any $m|n$ we have $\Delta(n) \leq \tau(n/m)\Delta(m)$ and we suppose that m is the product of the distinct prime factors of n and denote this by writing $n \in (m)$. Thus

$$S(x) \leq \sum_{m < x} |\mu(m)|\Delta(m) \sum \{ \tau(n/m) : n < x, n \in (m) \} \leq \sum_{m < x} |\mu(m)|\Delta(m) Q^2(x, m)$$

where Q is the function defined in Lemma 4. Let $h > 1$ and $k = 2h/(h-1)$. By Hölder's inequality,

$$\begin{aligned} S(x) &\leq \sum_{m < x} |\mu(m)|(\Delta(m))^{1/h}(\tau(m))^{2/k} Q^2(x, m) \\ &\leq (S_1(x))^{1/h} \left(\sum_{m < x} (\tau(m))^2 \right)^{1/k} \left(\sum_{m < x} |\mu(m)|Q^{2k}(x, m) \right)^{1/k} \\ &\ll_{\epsilon, h} x(\log x)^{(A_1 + \epsilon)/h + 3/k}, \end{aligned}$$

by Lemma 4 and the definition of A_1 . We let $h \rightarrow 1$ and obtain the result stated.

LEMMA 8. Let $\omega(n, t)$ denote the number of distinct prime factors of n which do not exceed t , and let z_0, y_0 be any fixed positive numbers. Then we have

$$\sum_{n < x} y^{\omega(n, t)} z^{\omega(n)} \ll x(\log x)^{z-1} (\log \min \{x, t\})^{y^2-z},$$

uniformly for $0 < y \leq y_0, 0 < z \leq z_0$ and all t .

This is a straightforward application of the very useful Theorem 2 of [3].

LEMMA 9. Let y_0, y_1, z_1 be fixed positive numbers, $y_1 \leq y_0$. Then we have

$$\sum_{\substack{n < x \\ \omega(n) = h}} y^{\omega(n, t)} \ll \frac{x \sqrt{(\log \log x)} (\log \log x + (y-1) \log \log \min \{x, t\})^h}{\log x \cdot h!}$$

uniformly for $y_1 \leq y \leq y_0, h \leq z_1 \log \log x$.

Proof. We restrict the sum in Lemma 8 to the integers n for which $\omega(n) = h$, and we set $z = h/(\log \log x + (y-1) \log \log \min \{x, t\})$. Hence $z \leq z_1/\min \{1, y_1\}$. This gives

$$\sum_{\substack{n < x \\ \omega(n) = h}} y^{\omega(n, t)} \ll \frac{x (\log \log x + (y-1) \log \log \min \{x, t\})^h}{\log x (h/e)^h}$$

and the result follows by Stirling's formula.

LEMMA 10. Let $K(x, \delta)$ denote the set of integers $n < x$ such that for every $t < x$ we have

$$\omega(n, t) - \omega(n) \frac{\log \log t}{\log \log x} < \delta \log \log x.$$

Then we have

$$\sum_{\substack{n \in K(x, \delta) \\ \omega(n) = k}} |\mu(n)| U(n, 0) < \frac{x(\log \log x)^{k+3/2}}{k!(\log x)^{1-\delta \log 3}} \left(1 + \frac{3^k}{\log x} \right),$$

uniformly for $k \leq v_0 \log \log x$, where v_0 is any fixed positive number.

Proof. Let $y \in (0, 1)$ be fixed and let $n \in K(x, \delta)$. Then

$$U(n, 0) \leq \sum'_{dd'|n} 1 \leq 2(\log x)^{-\delta \log y} \sum''_{dd'|n} y^{\omega(n, d) - v \log \log d}$$

where $v = k/\log \log x$ and \sum'' denotes summation restricted by the conditions that $d < d' < ed$, $(d, d') = 1$; it is convenient in what follows to assume that $d' > d$ and the factor 2 on the right takes this into account. Hence

$$\begin{aligned} \sum_{\substack{n \in K(x, \delta) \\ \omega(n) = k}} |\mu(n)| U(n, 0) &\leq 2(\log x)^{-\delta \log y} \sum_{\substack{n < x \\ \omega(n) = k}} |\mu(n)| \sum'_{dd'|n} y^{\omega(n, d) - v \log \log d} \\ &\leq 2(\log x)^{-\delta \log y} \sum'_{\substack{dd' < x \\ \omega(dd') \leq k}} y^{\omega(dd') - 1 - v \log \log d} \sum_{\substack{m < x/dd' \\ \omega(m) = k - \omega(dd')}} y^{\omega(m, d)}. \end{aligned}$$

The extra factor y^{-1} takes account of the possibility that d' has a prime factor greater than d ; it cannot have two. We split this sum into two parts Z_1, Z_2 according as $d < x^{1/3}$ or not. For brevity, we only give the details concerning Z_1 ; Z_2 is easier because we then have $d \geq x/d^2$ and so $\omega(m, d) = \omega(m)$ in the inner sum. Put $k - \omega(dd') = h$. By Lemma 8 (with $y_1 = 1/3$, our eventual choice for y), we have

$$Z_1 < (\log x)^{-\delta \log y} \sqrt{(\log \log x)} \sum_{a+b+h=k} y^{a+b} Z(a, b, h)$$

where

$$Z(a, b, h) = \sum_{\substack{d < x^{1/3} \\ \omega(d) = a}} \sum_{\substack{d < d' < ed \\ \omega(d') = b}} y^{-v \log \log d} \frac{x(\log \log x + (y-1) \log \log d)^h}{h! d^2 \log x}.$$

It is well known that there exist absolute constants C, D such that

$$\text{card} \{n : n < x, \omega(n) = r\} \leq \frac{Cx}{\log x} \frac{(\log \log x + D)^{r-1}}{(r-1)!}$$

for all positive x and positive integers r . Hence

$$\text{card} \{n : n < x, \omega(n) = r\} < \frac{x}{\log x} \frac{(\log \log x)^r}{r!}$$

uniformly for $r \leq r_0 \log \log x$, with any fixed r_0 . We apply this to the inner sum over d' above, and we deduce that

$$\sum_{b+h=k-a} y^b Z(a, b, h) \ll \frac{x}{\log x} \sum_{\substack{d < x^{1/3} \\ \omega(d) = a}} \frac{y^{-v \log \log d} (\log \log x + (2y-1) \log \log d)^{k-a}}{(k-a)! d \log d}.$$

We write this as a Stieltjes integral and integrate by parts. Thus the left hand side above is

$$\ll \frac{x}{\log x} \left\{ \frac{(\log \log x)^a (2y \log \log x)^{k-a}}{a!(k-a)! (\log x)^{2+v \log y}} + \int_1^{x^{1/3}} \frac{(\log \log t)^a (\log \log x + (2y-1) \log \log t)^{k-a}}{a!(k-a)! t (\log t)^{2+v \log y}} dt \right\}.$$

We multiply by y^a and sum over a . Choosing $y = 1/3$, and recalling that $v = k/\log x$, we obtain that

$$\sum_{a+b+h=k} y^{a+b} Z(a, b, h) \ll \frac{x (\log \log x)^{k+1}}{k! \log x} \left\{ 1 + \frac{3^k}{\log x} \right\}.$$

(We only need the extra factor $\log \log x$ if $v \log 3 = 1$.) Together with the corresponding upper bound for Z_2 , this gives the result stated.

LEMMA 11. *Uniformly for $0 < y \leq \log 2/\log 3$, we have that*

$$\sum_{\substack{n < x \\ n \in K(x, \delta)}} |\mu(n)| \Delta(n) y^{\omega(n)} \ll x (\log x)^{y-1+\delta \log 2} \log \log x.$$

Proof. If $\omega(n) > 10 \log \log x$ we use the trivial estimate $\Delta(n) \leq \tau(n)$. Otherwise put $\omega(n) = k$, and apply Lemmas 2 and 9. By Hölder's inequality, with exponents $\beta, \beta/\beta - 1$, we have

$$\begin{aligned} \sum_{\substack{n \in K(x, \delta) \\ \omega(n) = k}} |\mu(n)| \Delta(n) &\leq \left(3 \sum_{\substack{n \in K(x, \delta) \\ \omega(n) = k}} |\mu(n)| U(n, 0) \right)^{1/\beta} \left(\sum_{\substack{n < x \\ \omega(n) = k}} 1 \right)^{1-1/\beta} \\ &\ll \frac{x (\log \log x)^{k+1}}{k! \log x} \left\{ 1 + \frac{2^k}{\log^{1/\beta} x} \right\} (\log x)^{\delta \log 2} \end{aligned}$$

(noting that $3/2\beta < 1$). We multiply by y^k and sum over k , and for the range of values of y given, the first term on the right dominates. This completes the proof.

LEMMA 12. *Let $\lambda(v) = v - v \log v - 1$. Then we have*

$$\text{card} \{n < x : n \notin K(x, \delta), \omega(n) = v \log \log x\} \ll x (\log x)^{2\lambda(v+\delta) + 2\lambda(v-\delta) - 3\lambda(v)}$$

uniformly in any fixed range $3\delta \leq v \leq v_0$.

Without the condition that $n \notin K(x, \delta)$, the correct exponent on the right would be $\lambda(v)$. This would follow from Lemma 3 with $y = v$. Since λ is concave the result above is an improvement on this.

Proof. Set $T_r = \exp \exp r^2$ for $r = 1, 2, 3, \dots$ and suppose that $\omega(n, t) - v \log \log t > \delta \log \log x$. There exists an r such that $T_r < t < T_{r+1}$, so that

$$\omega(n, T_{r+1}) - v \log \log T_{r+1} > \delta \log \log x - (2r + 1)v > \delta \log \log x - 3v\sqrt{\log \log x}$$

since $r < \sqrt{\log \log x}$. Since v is bounded, the right hand side is $(\delta + o(1)) \log \log x$. Since λ is continuous and there are $< \sqrt{\log \log x}$ points T_r , it will be sufficient to prove that at each of them we have

$$\begin{aligned} \text{card} \{n : n < x, \omega(n) = v \log \log x, \omega(n, T_r) - v \log \log T_r > \delta \log \log x\} \\ << x(\log x)^{2\lambda(v+\delta) + 2\lambda(v-\delta) - 3v - \varepsilon} \end{aligned}$$

where $\varepsilon = \varepsilon(v, \delta) > 0$.

We set $s = (\log \log T_r) / (\log \log x) \leq (v - \delta) / v$ (since $\omega(n, T_r) \leq \omega(n)$). By Lemma 8 we have that

$$\sum_{n < x} y^{\omega(n, T_r)} z^{\omega(n)} << x(\log x)^{z-1 + sz(y-1)}$$

uniformly in any fixed ranges $z \leq z_0, y \leq y_0$. We suppose that $y \geq 1$, so that the cardinality in question is

$$<< x(\log x)^{z-1 + sz(y-1) - v \log z - (vs + \delta) \log y} << x(\log x)^{\lambda(v) + v \log(1 + s(y-1)) - (vs + \delta) \log y}$$

if we set $z = v / (1 + s(y-1))$. Since $y \geq 1$, we have $z \leq v_0$ and so we may take $z_0 = v_0$. When we minimize the exponent on the right by varying y , there is a slight complication inasmuch as the optimal y is unbounded as a function of s . We therefore select a large, fixed y_0 and set $y = \min(y_0, y(s))$ where $y(s) = (vs + \delta)(1 - s) / s(v - \delta - sv) > 1$. We have $y(s) \leq y_0$ in the range $s_0 \leq s \leq s_1$, where

$$s_0, s_1 = \frac{v - \delta}{2v} \pm \sqrt{\left(\left(\frac{v - \delta}{2v}\right)^2 - \frac{\delta}{v(y_0 - 1)}\right)}.$$

We have

$$s_0 \leq \frac{2\delta}{v - \delta} \frac{1}{y_0 - 1} \leq \frac{1}{y_0 - 1}.$$

Next, we show that with this choice of y , we have

$$v \log(1 + s(y - 1)) - (vs + \delta) \log y < 2\lambda(v + \delta) + 2\lambda(v - \delta) - 4\lambda(v)$$

for all s . We therefore have to maximize the function of s on the left in the range $0 \leq s \leq (v - \delta) / v$, and we prove first that we need only consider the central range $s_0 \leq s \leq s_1$. Outside this range, $y = y_0$ and so

$$\frac{d}{ds} \left(v \log(1 + s(y - 1)) - (vs + \delta) \log y \right) = \frac{v(y_0 - 1)}{1 + s(y_0 - 1)} - v \log y_0$$

and the derivative is positive if and only if

$$s < \frac{1}{\log y_0} - \frac{1}{y_0 - 1}.$$

Provided that $y_0 - 1 \geq 2 \log y_0$, as we may assume, this is the case for $0 \leq s < s_0$, and so the maximum does not lie in this range. Next, for any $y_0 > 1$ we have $(\log y_0)^{-1} - (y_0 - 1)^{-1} < 1/2$. Hence the derivative above is negative for $s \geq 1/2$, and we have

$$s_1 = \frac{v - \delta}{v} - s_0 \geq \frac{2}{3} - \frac{1}{y_0 - 1} \geq \frac{1}{2}$$

for suitably large y_0 . Hence the maximum we seek lies in the interval $[s_0, s_1]$. For such s , $y = y(s)$, and so

$$\begin{aligned} &v \log(1 + s(y - 1)) - (vs + \delta) \log y \\ &= v \log\left(\frac{v(1 - s)}{v - \delta - sv}\right) - (vs + \delta) \log\left(\frac{(vs + \delta)(1 - s)}{s(v - \delta - sv)}\right) \\ &= (v - \delta - sv) \log\left(\frac{1 - s}{v - \delta - sv}\right) + v \log v - (vs + \delta) \log\left(\frac{vs + \delta}{s}\right) \\ &= \int_0^\delta \log\left(\frac{s(v - \eta - sv)}{(1 - s)(vs + \eta)}\right) d\eta. \end{aligned}$$

We obtain an upper bound by maximizing the integrand as a function of s . The maximum occurs when $s = (v - \eta)/2v$ and so the integral above is less than

$$2 \int_0^\delta \log\left(\frac{v - \eta}{v + \eta}\right) d\eta = 2\lambda(v + \delta) + 2\lambda(v - \delta) - 4\lambda(v).$$

There is a strict inequality at this point because the s which maximizes the integrand varies with η . We denote by $\varepsilon(v, \delta)$ the difference between the true maximum and the function on the right above. This gives the result stated.

LEMMA 13. *We have that*

$$\sum_{\substack{n < x \\ n \notin K(x, \delta)}} z^{\omega(n)} \ll x(\log x)^{\mu(z, \delta)} \log \log x$$

uniformly for $0 \leq \delta \leq \frac{1}{10}$, $\frac{1}{2} \leq z \leq 1$, where

$$\mu(z, \delta) = \max_v \{v \log z + 2\lambda(v + \delta) + 2\lambda(v - \delta) - 3\lambda(v)\}.$$

Proof. The function $\lambda(v + \delta) + \lambda(v - \delta) - 2\lambda(v)$ increases with v : it is greater than

or equal to $\lambda(2\delta) + \lambda(0) - 2\lambda(\delta) = -\delta \log 4$. It follows that

$$\mu(z, \delta) \geq \max_v \{v \log z + \lambda(v)\} - 2\delta \log 4 \geq z - 1 - \delta \log 16.$$

The exponent of $\log x$ therefore exceeds -1 . We split the integers $n < x$ into three classes according as $\omega(n) > 4 \log \log x$, $4 \log \log x \geq \omega(n) \geq 3\delta \log \log x$, $3\delta \log \log x > \omega(n)$. The contribution from the first class is

$$\leq \sum_{\substack{n < x \\ \omega(n) > 4 \log \log x}} 1 \leq 4^{-4 \log \log x} \sum_{n < x} 4^{\omega(n)} \ll x(\log x)^{3-4 \log 4} \ll x(\log x)^{-1},$$

using Lemma 3. This is negligible. In the second class, put $\omega(n) = k = v \log \log x$, $3\delta \leq v \leq 4$. There are $\ll \log \log x$ different k and for each one, applying Lemma 12, the contribution to our sum is

$$\ll z^k x(\log x)^{2\lambda(v+\delta) + 2\lambda(v-\delta) - 3\lambda(v)}.$$

The exponent of $\log x$ is $v \log z + 2\lambda(v + \delta) + 2\lambda(v - \delta) - 3\lambda(v) \leq \mu(x, \delta)$. It remains to deal with the third class where we have $\omega(n) < 3\delta \log \log x$. We have

$$\sum_{\substack{n < x \\ \omega(n) < 3\delta \log \log x}} z^{\omega(n)} \leq y^{-3\delta \log \log x} \sum_{n < x} (yz)^{\omega(n)} \ll x(\log x)^{yz-1-3\delta \log y}$$

for any $y \leq 1$. We put $y = 3\delta/z$ so that the exponent of $\log x$ is $3\delta - 1 - 3\delta \log(3\delta/z)$. We have to show that this does not exceed $\mu(z, \delta)$. We define (for $0 < t < 1$),

$$h(t) = 1 - 2(1+t) \log(1+t) - 2 \log(1-t) \log(1-t) - 3t + 3t \log 3t.$$

Thus

$$h'(t) = \log \left\{ \left(\frac{1-t}{1+t} \right)^2 \cdot 27t^3 \right\} < 0 \quad \text{for } t < \frac{1}{3},$$

$$h\left(\frac{1}{3}\right) = \frac{1}{3} \{2 - \log(3^9 4^{14} / 5^{17})\} > 0.$$

Thus h is positive for $0 \leq t \leq 1/5$. But

$$\begin{aligned} \mu(z, \delta) &\geq z \log z + 2\lambda(z + \delta) + 2\lambda(z - \delta) - 3\lambda(z) \\ &\geq z - 1 + 4z \log z - 2(z + \delta) \log(z + \delta) - 2(z - \delta) \log(z - \delta). \end{aligned}$$

Hence $\mu(z, \delta) - \{3\delta - 1 - 3\delta \log(3\delta/z)\} \geq zh(\delta/z) \geq 0$. Thus the contribution of the third class integers is negligible. This completes the proof.

Proof of Theorem 2. By Lemma 7, it is sufficient to obtain the upper bound for A_1 . Clearly all our upper bounds, and the iteration involved in the proof of Theorem 1, are equally valid for squarefree numbers: we therefore restrict ourselves to these numbers, defining $f(y)$ as before. We improve our upper bound for $f(y)$ in the range $1/2 < y < 3/5$. Now

$$2f(1) \leq \min \{f(y) + 2y - 2 + (1/y) : 1/2 < y < 2\},$$

and when we substitute $f(y) = y^2 - 3/4$ the minimum occurs at the root of $2y^2(y + 1) = 1$. This root belongs to $(1/2, 3/5)$. Hence a smaller $f(y)$ in this range will give a better estimate for $f(1)$. We do not give any numerical details.

We split the squarefree numbers $n < x$ into two classes according as they lie in $K(x, \delta)$ or not. Here $\delta \in (0, 1/10)$ is at our disposal. By Lemma 11, for the integers in the first class we have

$$\sum_{n \in K(x, \delta)} |\mu(n)| \Delta(n) y^{\omega(n)} \ll x(\log x)^{y-1+\delta \log 2} \log \log x.$$

For the integers in the second class we have

$$\sum_{\substack{n < x \\ n \notin K(x, \delta)}} \Delta(n) y^{\omega(n)} \leq \left(\sum_{n < x} \Delta^2(n) 2^{-\omega(n)} \right)^{1/2} \left(\sum_{\substack{n < x \\ n \notin K(x, \delta)}} (2y^2)^{\omega(n)} \right)^{1/2}.$$

The first sum on the right was estimated during the course of the proof of Theorem 1. As n is squarefree we have $\omega(n) = \Omega(n)$ and so it does not exceed

$$T(x, \frac{1}{2}) \ll x(\log x)^{-1/2} (\log \log x)^2.$$

The second sum on the right is that which appears in Lemma 13. Putting these results together we deduce that

$$\sum_{\substack{n < x \\ n \notin K(x, \delta)}} \Delta(n) y^{\omega(n)} \ll x(\log x)^{\mu(2y^2, \delta)/2 - 1/4} (\log \log x)^{3/2}.$$

For all z , and $\delta > 0$,

$$\mu(z, \delta) < \max_v \{v \log z + \lambda(v)\} = z - 1$$

and so for $\delta > 0$, we have $\mu(2y^2, \delta)/2 - 1/4 < y^2 - 3/4$. For each $y < \frac{1}{2}$ there exists $\delta > 0$ for which $y - 1 + \delta \log 2 < y^2 - 3/4$; indeed we may choose $\delta = \delta(y)$ optimally by making it the root of the transcendental equation

$$y - 1 + \delta \log 2 = \mu(2y^2, \delta)/2 - 1/4.$$

We now have

$$f(y) \leq y - 1 + \delta(y) \log 2, \quad (1/2 < y < 3/5),$$

$$2f(1) < \min \{3y - 3 + \delta(y) \log 2 + (1/y) : 1/2 < y < 3/2\} < A_0.$$

This completes the proof.

References

1. P. Erdős, "On the density of some sequences of integers", *Bull. Amer. Math. Soc.*, 54 (1948), 685-692.
2. P. Erdős and G. Tenenbaum, "Sur la structure de la suite des diviseurs d'un entier", *Ann. Inst. Fourier (Grenoble)*, 31 (1981), 17-37.
3. H. Halberstam and H.-E. Richert, "On a result of R. R. Hall", *J. Number Theory*, 11 (1979), 76-89.
4. R. R. Hall, "The propinquity of divisors", *J. London Math. Soc.* (2), 19 (1979), 35-40.

5. R. R. Hall, "On the probability that n and $f(n)$ are relatively prime, III", *Acta. Arith.*, 20 (1972), 226–289.
6. R. R. Hall and G. Tenenbaum, "Sur la proximité des diviseurs", *Recent progress in analytic number theory* (eds. H. Halberstam and C. Hooley, Academic Press, London, 1981), pp.103–113.
7. C. Hooley, "On a new technique and its applications to the theory of numbers", *Proc. London Math. Soc.* (3), 38 (1979), 115–151.
8. D. R. Woodall, "A problem on cubes", *Mathematika*, 24 (1977), 60–62.

Department of Mathematics,
University of York,
York YO1 5DD.

UER de Mathématiques,
Université de Bordeaux I,
F 33405 Talence Cedex,
France.