J. Austral. Math. Soc. (*Series A*) **26** (1978), 487–500

A NEW DEFINITION OF THE DENSITY OF AN INTEGER SEQUENCE

R. R. HALL

(Received 12 April 1976; revised 19 July 1978)

Communicated by J. Pitman

Abstract

The divisor density of an integer sequence is defined: this measures the tendency of divisors of (almost all) integers to belong to the sequence. A proof of a conjecture of Erdös is given and this is linked to some previous conjectures of the author's concerning distribution (mod 1) of functions of the divisors.

Subject classification (*Amer. Math. Soc.* (*MOS*) *1970*): 10 L 10, 10 K 05.

1. Introduction

Let A be a strictly increasing sequence of positive integers. In this paper I introduce a new definition of the density of A which is appropriate for some divisor problems.

Let $\tau(n)$ denote the number of divisors of n and $\tau(n, A)$ the number of these divisors belonging to A.

DEFINITION. A possesses *divisor density* DA if $\tau(n, A) \sim DA \cdot \tau(n)$ on a sequence of integers n having asymptotic density 1.

REMARKS. Suppose that

$$\sum \left\{\frac{1}{d}: d \in A\right\} = \infty.$$

Then

$$\sum_{n \leqslant x} \tau(n, A) \sim x \sum_{d \leqslant x} \{d^{-1}: d \in A\}$$

and if A has logarithmic density δA the right-hand side is $\sim \delta A x \log x$, that is,

$$\sum_{n \leqslant x} \tau(n, A) \sim \delta A \sum_{n \leqslant x} \tau(n).$$

This might suggest that $DA = \delta A$; in fact this is quite misleading, and we shall show that there is no connection between divisor density and logarithmic density. More precisely, given $z \in [0, 1]$, $w \in [0, 1]$, we can construct A so that $\delta A = z$ while DA does not exist, and vice versa, or such that we have $\delta A = z$, $DA = w$.

It is not immediate, as in the case of asymptotic and logarithmic density, that for every $z \in [0, 1]$ there exists a sequence A such that $DA = z$. But we may deduce from Hall (1974b), Theorem 1 that the sequence

$$A = \{d \colon \log d \leqslant z (\mathrm{mod}\, 1)\} \tag{0}$$

has the required property. This particular A has $\delta A = z$, but fails to have asymptotic density.

Let $\nu(d)$ denote the number of distinct prime factors of d. For any fixed integers a and b ($b > 0$) consider the sequence $\{d \colon \nu(d) \equiv a (\mathrm{mod}\, b)\}$. The asymptotic density is defined, and is $1/b$; and we now show that in this case $DA = 1/b$. Notice that

$$\tau(n, A) = \frac{1}{b} \sum_{r=1}^{b} e^{-2\pi i a r/b} \sum_{d \mid n} e^{2\pi i \nu(d) r/b}. \tag{1}$$

Let m be the product of those prime factors p of n such that $p^2 \nmid n$. From (1), for $b \geqslant 2$,

$$\left| \tau(n, A) - \frac{1}{b} \tau(n) \right| \leqslant \tau(n/m) \left(2 \cos \frac{\pi}{b} \right)^{\nu(m)} = \tau(n) \left(\cos \frac{\pi}{b} \right)^{\nu(m)}$$

and the result follows from the fact that there exists a sequence of integers n of asymptotic density 1 on which $\nu(m) \to \infty$; indeed as a function of n, $\nu(m)$ has normal order $\log \log n$. This example leads to a simple construction of a sequence A with both divisor and asymptotic density equal to z. Let $z = \cdot \varepsilon_1 \varepsilon_2 \varepsilon_3 \ldots$ be expanded in binary form, and let

$$A_i = \{d \colon \nu(d) \equiv 2^{i-1} (\mathrm{mod}\, 2^i)\}.$$

Then

$$A = U\{A_i \colon \varepsilon_i = 1\}$$

has the property required.

Next, I would like to justify my earlier remark that divisor and logarithmic density are unconnected. Let $B(t) = \{d \colon \nu(d) < t \log \log d\}$. Then I shall show below that if $\frac{1}{2} < t < 1$, we have $DB = 1$, $\delta B = 0$. Suppose for a moment that we have proved this for some t. We can find A_1 such that δA_1 is either undefined or equal to z, also A_2 such that DA_2 is either undefined or equal to w: notice that DA_2 is

https://doi.org/10.1017/S1446788700012003 Published online by Cambridge University Press

undefined for any arithmetic progression other than Z^+ itself. Let $C = Z^+ \setminus B$ be the complement of B. Then $A = (A_1 \cap C) \cup (A_2 \cap B)$ has the required property, that is, δA and DA are whatever we like. Now $\delta B(t) = 0$ for $t < 1$ follows from the fact that for each $t < 1$ there exists $\eta = \eta(t) > 0$ such that

$$\operatorname{card}\{B(t) \cap [0, x]\} = O(x/(\log x)^{\eta}),$$

and hence

$$\sum (1/d\colon d \in B(t)) < \infty.$$

Now we show that $DB(t) = 1$ for $t > \frac{1}{2}$. Let $n = p_1^{\alpha_1} p_2^{\alpha_2} \dots p_r^{\alpha_r}$. Then

$$\sum_{d|n} (\nu(d) - \tfrac{1}{2}\nu(n))^2 = \tau(n)\left\{\left(\sum_{i=1}^{r} \frac{\alpha_i}{1+\alpha_i} - \tfrac{1}{2}\nu(n)\right)^2 + \sum_{i=1}^{r} \frac{\alpha_i}{(1+\alpha_i)^2}\right\} \tag{2}$$

$$\leqslant \tfrac{1}{4}\tau(n)\{(\omega(n) - \nu(n))^2 + \nu(n)\},$$

where $\omega(n) = \sum \alpha_i$. The normal order of $\nu(n)$ is $\log\log n$; moreover it is plain that $\omega(n) - \nu(n) < \sqrt{(\log\log n)}$ for almost all n, since the average order of $\omega(n) - \nu(n)$ is an absolute constant. Hence for almost all n, the right-hand side of (2) does not exceed $\tau(n)\log\log n$; moreover $\nu(n) < (1+\eta)\log\log n$ for any fixed, positive η. Hence $\nu(d) < (\frac{1}{2}+\eta)\log\log n$ for all but $O(\tau(n)/\log\log n)$ divisors of d of n. Let η be fixed so that $(\frac{1}{2}+\eta)/(1-\eta) = t$. When $t > \frac{1}{2}$, η is positive. The result will follow if we show that for almost all n, and all but $o(\tau(n))$ divisors d of n, we have $\log\log d > (1-\eta)\log\log n$. Let $\nu(n, \varepsilon)$ denote the number of distinct prime factors p of n such that $\log\log p > (1-\varepsilon)\log\log n$. By the familiar variance method due to Turàn, we can show that for fixed ε, $0 < \varepsilon < 1$, $\nu(n, \varepsilon)$ has normal order $\varepsilon \log\log n$. Set $\varepsilon = \eta$. Then for almost all n, $\nu(n, \eta) > \frac{1}{2}\eta\log\log n$, and if d is divisible by any prime counted by $\nu(n, \eta)$, plainly $\log\log d > (1-\eta)\log\log n$. The number of divisors not so divisible is $\leqslant 2^{-\nu(n,\eta)}\tau(n) = o(\tau(n))$, the desired result.

An interesting special case arises when A consists of relatively long blocks of consecutive integers. Let $\{b_j\}$ be an increasing, unbounded sequence of positive reals and $A = \{d\colon b_{2j} < d \leqslant b_{2j+1}$ for some $j\}$. Then the arithmetical properties (such as number of prime factors) of the numbers $d \in (b_{2i}, b_{2i+1}]$ should average out over the interval if this is long enough. Professor Erdös conjectured the truth of the following theorem which is one of the main results of this paper.

THEOREM 1. *Let $b_{j+1} > cb_j$ for fixed $c > 1$ and every j, and*

$$A = \{d\colon \exists j\colon b_{2j} < d \leqslant b_{2j+1}\}.$$

Let $\delta A = z$. Then $DA = z$.

I shall derive from this the following corollary.

https://doi.org/10.1017/S1446788700012003 Published online by Cambridge University Press

COROLLARY. *Let* $f\colon \mathbf{R}^{+}\to\mathbf{R}$ *be continuously differentiable and* $f'(x)\to 0$ *as* $x\to\infty$. *Let the sequence* $\{f(n)\colon n\in Z^{+}\}$ *be uniformly distributed* (mod 1). *Then for all* $z\in[0,1]$, *the sequence*

$$A' = \{d\colon f(\log d)\leqslant z (\operatorname{mod} 1)\}$$

has $DA' = z$.

For example, the sequences

$$A'' = \{d\colon (\log d)^{\alpha}\leqslant z (\operatorname{mod} 1)\}$$

and

$$A'' = \{d\colon (\log\log d)^{\beta}\leqslant z (\operatorname{mod} 1)\}$$

have $DA'' = z$, when $0<\alpha<1$, $\beta>1$ respectively. As explained below, it is probably sufficient that $\alpha>0$, but the condition $\beta>1$ is sharp.

We refer to these sequences of Erdös consisting of the blocks of integers $(b_{2j}, b_{2j+1}]\cap Z$ simply as block sequences. Thus (0) is a block sequence. Another way of describing a block sequence is to define an increasing, continuously differentiable function g such that $g(b_{2j}) = j$, $g(b_{2j+1}) = j+z$ for all j. Then

$$A = A(z) = \{d\colon g(d)\leqslant z (\operatorname{mod} 1)\}.$$

This gives a link with my paper, Hall (1976), where I defined g to be *uniformly distributed* if for almost all n, we have that

$$D_g(n,z) := \operatorname{card}\{d\colon d\,|\,n, g(d)\leqslant z (\operatorname{mod} 1)\}\sim z\tau(n)$$

uniformly for $0\leqslant z\leqslant 1$. Thus if g is *uniformly distributed* we have $DA(z) = z$ for all $z\in[0,1]$. The converse goes through except for the uniformity in z. An immediate consequence of the theorem proved in Hall (1976), and these remarks, is the following theorem.

THEOREM 2. *In the above circumstances, suppose that for* $u>u_0$, *we have*

(i) $ug'(u)$ *is monotonic*;

(ii) $\log^{-\gamma} u < |ug'(u)| < \log^{\gamma} u$, *where* $\gamma<\log(4/3)$.

Then $DA(z) = z$ *for all* $z\in[0,1]$.

For the sequences A'' of the first type mentioned above, this requires $|\alpha-1|<\log(4/3)$ and so extends the range to $0<\alpha<1+\log(4/3)$. Theorem 2 gives nothing if g varies too slowly, nevertheless I conjectured in Hall (1976) that $g(u) = (\log\log u)^{\beta}$ is uniformly distributed for $\beta>1$ and showed that this would be sharp. The same reasoning applies in the present case.

It is natural to ask for a necessary and sufficient condition on A for DA to exist, and equal z. I shall prove the following result which gives a necessary condition.

THEOREM 3. *If $DA = z$ then*

$$\sum_{d<x}{}' Q(d) \sim z \sum_{d<x} Q(d), \tag{3}$$

where the dash denotes that $d \in A$. $Q(d)$ is a multiplicative function defined on the prime powers by the formula

$$Q(p^\alpha) = p^{-\alpha}\left(\frac{1}{\alpha+1}+\frac{1}{(\alpha+2)p}+\dots\right)\left(1+\frac{1}{2p}+\frac{1}{3p^2}+\dots\right)^{-1}.$$

This is certainly not sufficient to imply $DA = z$, for it would ascribe a divisor density to the arithmetic progressions, the sequence of squarefree numbers and so on, which they do not possess. I conjecture that there is no sufficient condition for $DA = z$ purely in terms of the average order of some arithmetical function (multiplicative or otherwise) as in (3). Something like Erdös blocks really are needed in addition, although I imagine that their length could be significantly reduced.

2. Proofs of the theorems

We begin with Theorem 1. We assume that $z \geqslant \frac{1}{2}$, the other case follows by complements. We define

$$S(x) = \sum_{n\leqslant x}\left(\frac{\tau(n,A)}{\tau(n)}-z\right)^2 = S_2(x) - 2zS_1(x) + z^2[x],$$

and we have to show that under the conditions of the theorem, $S(x) = o(x)$. It is sufficient that

$$S_i(x) = \sum_{n\leqslant x}\left(\frac{\tau(n,A)}{\tau(n)}\right)^i \sim z^i x \quad (i = 1, 2),$$

and we only consider the case $i = 2$ as the other is similar and easier. Now

$$S_2 = S_2(x) = \sum_{\substack{d_1\leqslant x\\ [d_1,d_2]\leqslant x}}{}' \sum_{d_2\leqslant x}{}' \sum_{m\leqslant x/[d_1,d_2]} 1/\tau^2(m[d_1,d_2]),$$

where the dash denotes summation restricted to A. If g is multiplicative, then $g(md)/g(d)$ is multiplicative as a function of m for each fixed d. We have the formula (uniform for $y \geqslant 1$ and $d \in Z^+$),

$$\sum_{m\leqslant y}\frac{\tau^2(d)}{\tau^2(md)} = \frac{A_1 q(d)\, y}{(\log 2y)^{3/4}} + O\left(\frac{y}{(\log 2y)^{7/4}}\prod_{p|d}\left(1+\frac{1}{\sqrt{p}}\right)\right), \tag{4}$$

where

$$A_1 = \frac{1}{\Gamma(\frac{1}{4})}\prod_p\left(1-\frac{1}{p}\right)^{1/4}\left(1+\frac{1}{4p}+\frac{1}{9p}+\dots\right)$$

https://doi.org/10.1017/S1446788700012003 Published online by Cambridge University Press

and q is the multiplicative function generated by

$$q(p^\alpha) = \left(1+\frac{(\alpha+1)^2}{(\alpha+2)^2 p}+\frac{(\alpha+1)^2}{(\alpha+2)^2 p^2}+\dots\right)\left(1+\frac{1}{4p}+\frac{1}{9p^2}+\dots\right)^{-1}.$$

I will give several such formulae in this paper; for a proof of this one see the Appendix, the proofs of the others are similar. In each case the main terms may be derived from a theorem of Wirsing (1967); however, to get a uniform error term we use the contour integration method starting with Perron's formula.

We deduce from the above that

$$S_2 = A_1 \sum_{\substack{d_1\leqslant x\\ [d_1,d_2]\leqslant x}}{}' \sum_{d_2\leqslant x}{}' \frac{q([d_1,d_2])\,x}{[d_1,d_2]\,\tau^2([d_1,d_2])(\log 2x/[d_1,d_2])^{3/4}}$$

$$+O\left(\sum_{r\leqslant x}\frac{\lambda(r)\,x}{r\tau^2(r)}(\log 2x/r)^{-7/4}\prod_{p\mid r}\left(1+\frac{1}{\sqrt{p}}\right)\right),$$

where $\lambda(r)$ is the number of solutions of $[d_1,d_2]=r$. If $r=p_1^{\alpha_1}p_2^{\alpha_2}\dots p_s^{\alpha_s}$ then $\lambda(r)=\prod(1+2\alpha_i)\leqslant 3^{\alpha_1+\dots+\alpha_s}$. The error term is therefore

$$\ll \sum_{r\leqslant x}\frac{\lambda(r)}{\tau^2(r)}\prod_{p\mid r}\left(1+\frac{1}{\sqrt{p}}\right)\sum_{m\leqslant x/r}(\log 2m)^{-7/4}$$

$$\ll \sum_{m\leqslant x}(\log 2m)^{-7/4}\sum_{r\leqslant x/m}\frac{\lambda(r)}{\tau^2(r)}\prod_{p\mid r}\left(1+\frac{1}{\sqrt{p}}\right).$$

The inner sum can be estimated as in (1974a); I give more details in the Appendix. We find that the above is

$$\ll \sum_{m\leqslant x}\frac{x}{m}(\log 2m)^{-7/4}(\log 2x/m)^{-1/4}\ll x(\log x)^{-1/4}. \tag{5}$$

In what follows, we estimate similar error terms the same way. Next we write $d_1=d$, $d_2=dm/k$ where $k\mid d$ and $(m,k)=1$. Thus $[d_1,d_2]=dm$. We use two dashes to denote $dm/k\in A$, and we have

$$S_2+O(x(\log x)^{-1/4}) = A_1\sum_{d\leqslant x}{}'\frac{x}{d}\sum_{k\mid d}\sum_{\substack{m\leqslant x/d\\(m,k)=1}}{}''\frac{q(md)}{\tau^2(md)}(\log 2x/md)^{-3/4}. \tag{6}$$

We have the formula, uniform for $y\geqslant 1$, all d, and all k dividing d,

$$\sum_{\substack{m\leqslant y\\(m,k)=1}}\frac{q(md)\,\tau^2(d)}{q(d)\,\tau^2(md)} = \frac{A_2\,y\,q(d,k)}{(\log 2y)^{3/4}}+O\left(\frac{y}{(\log 2y)^{7/4}}\prod_{p\mid d}\left(1+\frac{1}{\sqrt{p}}\right)\right),$$

where

$$A_2 = \frac{1}{\Gamma(\frac14)}\prod_p\left(1-\frac{1}{p}\right)^{1/4}\left(1+\frac{q(p)}{4p}+\frac{q(p^2)}{9p^2}+\dots\right)$$

https://doi.org/10.1017/S1446788700012003 Published online by Cambridge University Press

and for each fixed k, $q(d,k)$ is a multiplicative function of d (defined whether $k|d$ or not). If $p\nmid k$ then

$$q(p^\alpha,k)=\left\{\sum_{\beta=0}^{\infty}\left(\frac{\alpha+1}{\alpha+\beta+1}\right)^2\frac{q(p^{\alpha+\beta})}{q(p^\alpha)p^\beta}\right\}\left\{\sum_{\beta=0}^{\infty}\frac{q(p^\beta)}{(\beta+1)^2p^\beta}\right\}^{-1}.$$

If p$|k$ the first factor on the right is omitted. By partial summation, we get for $u<v\leqslant x/d$,

$$\sum_{m+u+1}^{v}\frac{q(md)\,\tau^2(d)}{mq(d)\,\tau^2(md)}(\log 2x/md)^{-3/4}$$

$$=A_2\sum_{m=u+1}^{v}\frac{q(d,k)}{m}(\log 2m)^{-3/4}(\log 2x/md)^{-3/4}$$

$$+O\left(\left\{(\log 2u)^{-7/4}(\log 2x/du)^{-3/4}+(\log 2v)^{-7/4}(\log 2x/dv)^{-3/4}\right.\right.$$

$$\left.\left.+\sum_{m=u+1}^{v}\frac{1}{m}(\log 2m)^{-7/4}(\log 2x/md)^{-3/4}\right\}\prod_{p|d}\left(1+\frac{1}{\sqrt{p}}\right)\right).$$

We can now estimate the inner sum in (6) which we suppose multiplied by $\tau^2(d)/q(d)$. The condition $md/k\in A$ means that m is restricted to blocks $u<m\leqslant v$ as above, where $u=u_i=[kb_{2i}/d]$, $v=v_i=[kb_{2i+1}/d]$ for some i. Let r and s be such that $b_{2r-1}<d/k\leqslant b_{2r+1}$, $b_{2s}<x/k\leqslant b_{2s+2}$ (we choose $r=0$ if $d/k\leqslant b_1$, $s=0$ if $x/k\leqslant b_2$), and let us set

$$a=\max(b_{2r},d/k),\quad b=\min(b_{2s+1},x/k).$$

These variables are of course functions of d and k. Then we have from (6) and the above that

$$S_2+O(x(\log x)^{-1/4})$$

$$=A_1A_2\sum_{d\leqslant x}{}'\frac{xq(d)}{d\tau^2(d)}\sum_{k|d}q(d,k)\sum_{i=r}^{s}\sum_{m=u_i+1}^{v_i}\frac{1}{m}(\log 2m)^{-3/4}(\log 2x/md)^{-3/4}$$

$$+O\left(\sum_{d\leqslant x}\frac{xq(d)}{d\tau^2(d)}\sum_{k|d}\left\{\sum_{i=r}^{s}(\log 2u_i)^{-7/4}(\log 2x/dv_i)^{-3/4}\right.\right.$$

$$\left.\left.+\sum_{m\leqslant x/d}\frac{1}{m}(\log 2m)^{-7/4}(\log 2x/md)^{-3/4}\right\}\prod_{p|d}\left(1+\frac{1}{\sqrt{p}}\right)\right).$$

To estimate the first error term we use the fact that the sequences $\{u_i\}$, $\{v_i\}$ increase geometrically. Thus the error term complete is

$$\ll\sum_{d\leqslant x}\frac{xq(d)}{d\tau(d)}(\log 2x/d)^{-3/4}\prod_{p|d}\left(1+\frac{1}{\sqrt{p}}\right)\ll x(\log x)^{-1/4}.$$

https://doi.org/10.1017/S1446788700012003 Published online by Cambridge University Press

Next, we consider the inner sum in the main term. We may change this to the corresponding integral without serious error, and if we denote the characteristic function of the intervals $(b_{2i}, b_{2i+1}]$ by $X(t)$, and make the change of variable $m = kt/d$, we obtain

$$S_2 + O(x(\log x)^{-1/4}) = A_1 A_2 \sum_{d \leqslant x}' \frac{xq(d)}{d\tau^2(d)} \sum_{k|d} q(d,k) \int_a^b \frac{t^{-1} X(t)\, dt}{(\log 2kt/d)^{3/4} (\log 2x/kt)^{3/4}},$$

where a and b are defined as above and depend on d and k. Let us set

$$\alpha = a\left(\frac{4x}{d}\right)^{\eta}, \quad \beta = b\left(\frac{4x}{d}\right)^{-\eta},$$

where $\eta = \eta(x) \in (0, \frac{1}{2})$, and consider the error involved in changing the range of integration to (α, β), taking into account that $0 \leqslant X(t) \leqslant 1$. We write

$$(7) \qquad 2kt/d = (4x/d)^{\omega}, \quad t^{-1}\, dt = (\log 4x/d)\, d\omega$$

and the error in the integral is

$$\leqslant \left(\log \frac{4x}{d}\right)^{-\frac{1}{2}} \int_I \frac{d\omega}{\omega^{3/4}(1-\omega)^{3/4}},$$

where $I \subseteq [0, \eta] \cup [1-\eta, 1]$. This is $\ll \eta^{1/4} (\log 4x/d)^{-\frac{1}{2}}$. Next, let us integrate over (α, β) by parts, writing

$$X_1(t) = \int_0^t s^{-1} X(s)\, ds = x \log t + E(t).$$

The error involved in replacing $X_1(t)$ by $z \log t$, or equivalently $X(t)$ by z in the original integral, is

$$\leqslant \frac{|E(\alpha)| + |E(\beta)|}{\eta^{3/4}(1-\eta)^{3/4}(\log 4x/d)^{3/2}} + \int_\alpha^\beta |E(t)| \left| \frac{d}{dt} \left(\log \frac{2kt}{d}\right)^{-3/4} \left(\log \frac{2x}{kt}\right)^{-3/4} \right| dt.$$

Since

$$\alpha \geqslant \frac{d}{2k}\left(\frac{4x}{d}\right)^{\eta}, \quad \beta \leqslant \frac{2x}{k}\left(\frac{4x}{d}\right)^{\eta} = \frac{d}{2k}\left(\frac{4x}{d}\right)^{1-\eta},$$

and the function $(\log 2kt/d)^{-3/4}(\log 2x/kt)^{-3/4}$ is convex and positive, this does not exceed

$$\frac{4M(2x/k)}{\eta^{3/4}(1-\eta)^{3/4}(\log 4x/d)^{3/2}},$$

where $M(x) = \max\{|E(y)| : y \leqslant x\}$. Hence we have

$$\int_a^b \frac{t^{-1} X(t)\, dt}{(\log 2kt/d)^{3/4}(\log 2x/kt)^{3/4}} = z \int_\alpha^\beta \frac{t^{-1}\, dt}{(\log 2kt/d)^{3/4}(\log 2x/kt)^{3/4}}$$
$$+ O(\eta^{1/4}(\log 4x/d)^{-\frac{1}{2}} + M(2x)\, \eta^{-3/4}(\log 4x/d)^{-3/2}),$$

and we may change the range of integration back to (a, b) with the same error term.

Having done this, we make the $t = t(\omega)$ substitution (7) and find that the integral on the right is

$$z\left(\log\frac{4x}{d}\right)^{-\frac{1}{2}} \int_{\xi}^{\zeta} \frac{d\omega}{\omega^{3/4}(1-\omega)^{3/4}} = J \quad \text{(say)},$$

where

$$\xi = \left(\log\frac{4x}{d}\right)^{-1}\log\frac{2ka}{d} \leqslant \left(\log\frac{4x}{d}\right)^{-1}\left(\log 2 + \log^{+}\frac{b_{2r}}{d/k}\right),$$

$$1-\zeta = \left(\log\frac{4x}{d}\right)^{-1}\log\frac{2x}{bk} \leqslant \left(\log\frac{4x}{d}\right)^{-1}\left(\log 2 + \log^{+}\frac{x/k}{b_{2s+1}}\right),$$

where as usual $\log^{+} y = \max(0, \log y)$. Suppose for example that $b_{2s+1} < x/k$. By hypothesis

$$0 = \int_{b_{2s+1}}^{x/k} t^{-1} X(t)\, dt = z \log\left(\frac{x/k}{b_{2s+1}}\right) + O(M(x/k).$$

It is at this point that we use the fact that $z \geqslant \frac{1}{2}$. We deduce that

$$\max(\xi, 1-\zeta) \ll (\log 4x/d)^{-1} M(x)$$

and so

$$J = B(\tfrac{1}{4},\tfrac{1}{4})\, z\left(\log\frac{4x}{d}\right)^{-\frac{1}{2}} + O\left(M^{1/4}(x)\left(\log\frac{4x}{d}\right)^{-3/4}\right).$$

The right-hand side does not depend on k, accordingly we define the function $\rho(d)$ by the relation

$$\tau(d)\,\rho(d) = \sum_{k|d} q(d,k)$$

and, using $M(x) = o(\log x)$, we have

$$\begin{aligned} S_2 &+ O(x(\log x)^{-1/4}) \\ &= A_1 A_2 B(\tfrac{1}{4},\tfrac{1}{4}) \sum_{d\leqslant x}{}' \frac{zxq(d)\,\rho(d)}{d\tau(d)(\log 4x/d)^{\frac{1}{2}}} \\ &\quad + O\left\{\sum_{d\leqslant x} \frac{xq(d)}{d\tau(d)} (\eta^{1/4}(\log 4x/d)^{-\frac{1}{2}} + \eta^{-3/4} M^{1/4}(2x)(\log 4x/d)^{-3/4})\right\}. \end{aligned}$$

We proceed by partial summation. We have

$$\sum_{d\leqslant y} \frac{q(d)\,\rho(d)}{\tau(d)} = \frac{A_3 y}{(\log 2y)^{\frac{1}{2}}} + O(y(\log 2y)^{-3/2}),$$

where

$$A_3 = \frac{1}{\sqrt{\pi}} \prod_p \left(1-\frac{1}{p}\right)^{\frac{1}{2}} \left(1 + \frac{q(p)\,\rho(p)}{2p} + \frac{q(p^2)\,\rho(p^2)}{3p^2} + \dots\right),$$

https://doi.org/10.1017/S1446788700012003 Published online by Cambridge University Press

and by similar reasoning to the above we finally arrive at

$$S_2 = A_1 A_2 A_3 B(\tfrac{1}{4}, \tfrac{1}{4}) B(\tfrac{1}{2}, \tfrac{1}{2}) z^2 x + O\left(\eta^{1/4} x + \eta^{-3/4} x \left(\frac{M(2x)}{\log x}\right)^{1/4}\right).$$

The factor multiplying $z^2 x$ is

$$\prod_p \left(1-\frac{1}{p}\right)\left(1+\frac{1}{4p}+\dots\right)\left(1+\frac{q(p)}{4p}+\dots\right)\left(1+\frac{q(p)\,\rho(p)}{2p}+\dots\right)$$

and this collapses to 1. Moreover as $M(x) = o(\log x)$ we can choose $\eta(x) \to 0$ so slowly that the error term is $o(x)$. We get $S_2 \sim z^2 x$ as required.

Next, we prove the corollary. Let us assume $0 < x < 1$, as the other cases are similar, and choose $\xi < \min(z, 1-z)$. Let

$$A^+ = \{d : f([\log d]) \leqslant z + \xi \pmod 1\},$$

$$A^- = \{d : f([\log d]) \leqslant z - \xi \pmod 1\}.$$

These sequences have the required blocks since $[\log d]$ is constant for $e^n < d < e^{n+1}$. Next

$$\sum \left\{\frac{1}{d} : e^n < d < e^{n+1}\right\} = 1 + O(e^{-n})$$

and since $f(n)$ is uniformly distributed (mod 1) we deduce that A^+, A^- have logarithmic density $z+\xi$, $z-\xi$ respectively. By Theorem 1,

$$\tau(n, A^+) \sim (z+\xi)\,\tau(n),$$

$$\tau(n, A^-) \sim (z-\xi)\,\tau(n),$$

on a sequence of asymptotic density 1. Next, there exists a $d_0 = d_0(\xi, f)$ such that for $d > d_0$, we have $|f(\log d) - f([\log d])| < \xi$, since $f' \to 0$. Hence for all n,

$$\tau(n, A^-) - d_0 \leqslant \tau(n, A) \leqslant \tau(n, A^+) + d_0.$$

Now consider the integers $n \leqslant x$. We can let $\xi = \xi(x) \to 0$ so slowly that

$$d_0(\xi, f) < \surd(\log x).$$

For all but $o(x)$ integers $n \leqslant x$ we have $\tau(n) > (\log x)^{3/5}$ (by the normal order property of τ) and so we have

$$|\tau(n, A) - z\tau(n)| < \xi\tau(n) + (\tau(n))^{5/6} = o(\tau(n))$$

for all but $o(x)$ integers $n \leqslant x$. Hence $DA = z$.

It remains to prove Theorem 3. If $DA = z$, then

$$zx \sim \sum_{n\leqslant x} \frac{\tau(n,A)}{\tau(n)} \sim \sum_{d\leqslant x}' \frac{1}{\tau(d)} \sum_{m\leqslant x/d} \frac{\tau(d)}{\tau(md)} \sim A_4 x \sum_{d\leqslant x}' \frac{Q(d)}{(\log 2x/d)^{\frac{1}{2}}}$$

by similar arguments to those used in the proof of Theorem 1. The value of A_4 is unimportant. As before, the dash denotes $d \in A$. Let

$$K(x,A) = \sum_{d\leqslant x}' Q(d).$$

Then we have proved that

$$\int_0^x \frac{dK(y,A)}{(\log 2x/y)^{\frac{1}{2}}} \to z/A_4 \quad \text{as } x\to\infty,$$

or if we put $x = e^t$, $y = e^u$,

(8) $$\int_0^t \frac{dK(e^u,A)}{(\log 2+t-u)^{\frac{1}{2}}} \to z/A_4 \quad \text{as } t\to\infty.$$

Define the Laplace transforms

(9)
$$\tilde{K}(s,A) = \int_0^\infty e^{-su}\, dK(e^u,A) \quad (\operatorname{Re} s>0)$$
$$h(s) = \int_0^\infty \frac{e^{-su}\,du}{\sqrt{(\log 2+u)}} \quad (\operatorname{Re} s>0).$$

We take the transform of (8), and apply the convolution formula for Laplace transforms. This gives

$$\tilde{K}(s,A)\,h(s) \sim (z/A_4)\,s^{-1} \quad \text{as } s\to 0+.$$

But $h(s) \sim \Gamma(\frac{1}{2})s^{-\frac{1}{2}}$ as $s\to 0+$, so that $\tilde{K}(s,A) \sim (z/A_4\,\Gamma(\frac{1}{2}))\,s^{-\frac{1}{2}}$ as $s \to 0+$. We now apply the Hardy–Littlewood–Karamata theorem, which states that if $\mu(u)$ is non-decreasing and

$$\int_0^\infty e^{-su}\,d\mu(u) \sim \Gamma(\delta+1)\,s^{-\delta} \quad \text{as } s\to 0+$$

for some fixed positive δ then

$$\int_0^T d\mu(u) \sim T^\delta \quad \text{as } T\to\infty.$$

(This is Theorem 98 of Hardy (1948) when $\delta = 1$: the more general case is a straightforward modification.) From (9) and the above, we deduce that

$$K(e^u,A) \sim (2/A_4\,\pi)\,u^{\frac{1}{2}} \quad \text{as } u\to\infty,$$

https://doi.org/10.1017/S1446788700012003 Published online by Cambridge University Press

that is

$$\sum_{d\leqslant x}' Q(d) \sim \frac{2z}{A_{4\pi}}\sqrt{(\log x)}.$$

In the special case $A = Z^+$, this gives

$$\sum_{d\leqslant x} Q(d) \sim \frac{2}{A_{4\pi}}\sqrt{(\log x)}$$

and putting these together, we have the result stated. The argument is not reversible.

3. Appendix

In this section we give outline proofs of (4) and (5), which also serve as models for the proofs of similar results used elsewhere in the paper.

Let us begin with (4). By Lemma 3.12 of Titchmarsh (1951) we have

$$\sum_{m\leqslant y} \frac{\tau^2(d)}{\tau^2(md)} = \frac{1}{2\pi i}\int_{c-iT}^{c+iT} F(s)\frac{y^s}{s}\,ds + O\left(\frac{y}{T}\log y\right),$$

where $c = 1+1/\log y$ and

$$F(s) = \sum_{m=1}^{\infty} \frac{\tau^2(d)}{\tau^2(md)\,m^s} = \prod_p \left(1+\frac{(\alpha+1)^2}{(\alpha+2)^2 p^s}+\frac{(\alpha+1)^2}{(\alpha+3)^2 p^{2s}}+\dots\right);$$

on the right α depends on p, it is the highest power of p which divides d. We can write

$$F(s) = F_0(s)\,G_d(s)$$

where

$$F_0(s) = \prod_p \left(1+\frac{1}{4p^s}+\frac{1}{9p^s}+\dots\right) = \{\zeta(s)\}^{1/4} g(s) \quad \text{(say)}$$

and

$$G_d(s) = \prod_{p\mid d} \left(1+\frac{(\alpha+1)^2}{(\alpha+2)^2 p^s}+\dots\right)\left(1+\frac{1}{4p^s}+\dots\right)^{-1}.$$

These functions are initially defined for $\operatorname{Re} s > 1$, but may be analytically continued: indeed $g(s)$ is regular for $\operatorname{Re} s > \frac{1}{2}$, and $G_d(s)$ is regular for $\operatorname{Re} s > 0$. We have

$$|g(s)| \ll 1, \quad |Ge(s)| \ll \prod_{p\mid d}\left(1+\frac{1}{\sqrt{p}}\right)$$

uniformly for $\operatorname{Re} s \geqslant \frac{3}{4}$. Next, it is known that $\zeta(s)$ does not vanish in the region

$$\operatorname{Re} s \geqslant \sigma_0(t) = 1 - C_0/\max(1, \log|t|)$$

https://doi.org/10.1017/S1446788700012003 Published online by Cambridge University Press

(for some fixed $C_0>0$, we assume $C_0>\frac{1}{8}$). We move the line of integration to $\operatorname{Re} s=\sigma_0(t)$ except for a loop around the algebraic singularity of $F_0(s)$ at $s=1$. The integral along $\operatorname{Re} s=\sigma_0(t)$ and along $\{s\colon \sigma_0(t)<\operatorname{Re} s<c, t=\pm T\}$ is

$$\ll \{y^{1-C_0/\log T}\log T+y/T\}\prod_{p\mid d}\left(1+\frac{1}{\sqrt{p}}\right)$$

and $T=\exp(\sqrt{\log y})$ gives a much better estimate than the error term stated in (4); so all this is negligible.

In the disc $|s-1|<\frac{1}{2}$, $F(s)(s-1)^{1/4}$ is regular and

$$\lim_{s\to 1} F(s)(s-1)^{1/4}=g(1)\,G_d(1)=\Gamma(\tfrac{1}{4})A_1 q(d).$$

Thus Cauchy's integral formula gives

$$\left|\frac{F(s)}{s}\frac{\Gamma(\frac{1}{4})A_1 q(d)}{(s-)^{1/4}}\right|\ll |s-1|^{3/4}\prod_{p\mid d}\left(1+\frac{1}{\sqrt{p}}\right)$$

uniformly for $|s-1|<\frac{1}{3}$. We use this approximation on the loop: putting $r=1-s$, the error involved is

$$\ll y\int_0^\infty y^{-r}r^{3/4}\prod_{p\mid d}\left(1+\frac{1}{\sqrt{p}}\right)dr\ll\frac{y}{(\log 2y)^{7/4}}\prod_{p\mid d}\left(1+\frac{1}{\sqrt{p}}\right)$$

for $y>2$ (for smaller y, our result is trivial). We now extend the loop to $s=-\infty$ with a further small error and use Hankel's integral formula for $1/\Gamma(z)$.

Next, we prove (5). Put $\beta=(\log 3)/(\log 2)$. Then $(1+x)^\beta\geqslant 1+2x$ for $x\geqslant 1$, and so $\lambda(r)\leqslant\{\tau(r)\}^\beta$. Since $\beta<2$, it readily follows that the multiplicative function

$$h(r)=\frac{\lambda(r)}{\tau^2(r)}\prod_{p\mid r}\left(1+\frac{1}{\sqrt{p}}\right)$$

is bounded. My result (Hall (1974a)) gives an estimate for sums of multiplicative functions such that $0\leqslant h(r)\leqslant 1$. This is easily modified to deal with bounded $h(r)$; in fact in a forthcoming paper, Halberstam and Richert give a much wider generalization. We find that

$$\sum_{r\leqslant x} h(r)\leqslant\left(1+O\left(\frac{\log\log x}{\log x}\right)\right)\frac{x}{\log x}\prod_p\left(1+\frac{h(p)}{p}+\frac{h(p^2)}{p^2}+\dots\right)$$

$$\ll x(\log x)^{-1/4}$$

in the present application.

https://doi.org/10.1017/S1446788700012003 Published online by Cambridge University Press

References

R. R. Hall (1974a), 'Halving an estimate obtained from Selberg's upper bound method', *Acta Arith.* **25**, 347–351.
R. R. Hall (1974b), 'The divisors of integers 1', *Acta Arith.* **26**, 41–46.
R. R. Hall (1976), 'The distribution of $f(d)$ (mod 1)', *Acta Arith.* **31**, 91–97.
G. H. Hardy (1948), *Divergent series* (Oxford).
E. C. Titchmarsh (1951), *The theory of the Riemann Zeta function* (Oxford).
E. Wirsing (1967), 'Das asymptotische Verhalten von Summen über multiplikative Functionen II', *Acta Math. Acad. Sci. Hung.* **18**, 411–467.

Department of Mathematics
University of York
Heslington, York YO1 5DD
England