

ROUGH INTEGERS WITH A DIVISOR IN A GIVEN INTERVAL

KEVIN FORD

(Received 19 February 2019; accepted 24 September 2019; first published online 8 January 2020)

Communicated by I. Shparlinski

Abstract

We determine, up to multiplicative constants, the number of integers $n \leq x$ that have a divisor in $(y, 2y]$ and no prime factor $\leq w$. Our estimate is uniform in x, y, w . We apply this to determine the order of the number of distinct integers in the $N \times N$ multiplication table, which are free of prime factors $\leq w$, and the number of distinct fractions of the form $(a_1 a_2)/(b_1 b_2)$ with $1 \leq a_1 \leq b_1 \leq N$ and $1 \leq a_2 \leq b_2 \leq N$.

2010 *Mathematics subject classification*: primary 11N25; secondary 62G30.

Keywords and phrases: divisors, Farey fractions.

1. Introduction

In the paper [5], the author established the order of growth of $H(x, y, z)$, the number of integers $n \leq x$, which have a divisor in the interval $(y, z]$, for all x, y, z . An important special case is

$$H(x, y, 2y) \asymp \frac{x}{(\log y)^\mathcal{E} (\log_2 y)^{3/2}} \quad (3 \leq y \leq \sqrt{x}), \quad (1-1)$$

where

$$\mathcal{E} = 1 - \frac{1 + \log_2 2}{\log 2} = 0.086071332 \dots$$

(The notation $\log_2 x$ stands for $\log \log x$.) A shorter, more direct proof of the order of magnitude bounds in the special case (1-1) is given in [6]. More on the history of estimations of $H(x, y, z)$, further applications and references may be found in [5].

A number of recent applications have required similar bounds, but where the underlying set of integers n is restricted to a special set, for example, the set of shifted primes ([5, Theorem 6,7], [9]) or the values of a polynomial [1, 4, 7, 12, 13]. More generally, we define

$$H(x, y, z; \mathcal{A}) = |\{n \leq x, n \in \mathcal{A} : d \mid n \text{ for some } d \in (y, z]\}|.$$

Research supported by National Science Foundation grant DMS-1802139.

© 2020 Australian Mathematical Publishing Association Inc.

Another natural set to consider is \mathcal{R}_w , the set of integers with no prime factor $p \leq w$; called *w-rough numbers* by some authors. Here we determine the exact order of growth of $H(x, y, 2y; \mathcal{R}_w)$ for all x, y, w ; the more general quantity $H(x, y, z; \mathcal{R}_w)$ can be estimated by similar methods, although there are many cases depending on the relative size of the parameters w, x, y, z .

THEOREM 1.1. *Suppose that $4 \leq y \leq \sqrt{x}$, $4 \leq w \leq y/8$ and write $\delta = \log_2 w / \log_2 y$.*

(i) *When $1 - 1/\log 4 \leq \delta \leq 1$,*

$$H(x, y, 2y; \mathcal{R}_w) - H(x/2, y, 2y; \mathcal{R}_w) \gg \frac{x}{\log^2 w} \gg H(x, y, 2y; \mathcal{R}_w).$$

(ii) *When $0 \leq \delta < 1 - 1/\log 4$,*

$$\begin{aligned} H(x, y, 2y; \mathcal{R}_w) - H(x/2, y, 2y; \mathcal{R}_w) &\gg x\delta B(w, y)(\log y)^{-\mathcal{E}+(\log(1-\delta))/\log 2} \\ &\gg H(x, y, 2y; \mathcal{R}_w), \end{aligned}$$

where

$$B(w, y) = \min(1, (\log_2 y)^{-1/2}((1 - \delta) \log 4 - 1)^{-1}).$$

REMARK 1.2. Some special cases are worth noting. From Theorem 1.1 we have, for each fixed $\varepsilon > 0$,

$$\begin{aligned} &x\delta B(w, y)(\log y)^{-\mathcal{E}+(\log(1-\delta))/\log 2} \\ &\asymp_{\varepsilon} \begin{cases} \frac{x \log_2 w}{(\log_2 y)^{3/2}} (\log y)^{-\mathcal{E}+(\log(1-\delta))/\log 2} & \left(\delta \leq 1 - \frac{1}{\log 4} - \varepsilon \right) \\ \frac{x \log_2 w}{(\log y)^{\mathcal{E}}(\log w)^{1/\log 2}(\log_2 y)^{3/2}} & (\log_2 w \leq \sqrt{\log_2 y}). \end{cases} \end{aligned}$$

REMARK 1.3. When $y > \sqrt{x}$, one can obtain similar results by using the duality $d | n \iff (n/d) | n$. That is, if $x/2 < n \leq x$, then $d | n$ with $y < d \leq 2y$ is equivalent to $d' | n$ with $d' \asymp x/y$.

We illustrate the utility of Theorem 1.1 with two applications. The first is related to the well-known multiplication table problem of Erdős [2, 3], which asks for estimates on the number, $M(N)$, of distinct integers in an $N \times N$ multiplication table. In [5] the author proved, using (1-1), that

$$M(N) \asymp \frac{N^2}{(\log N)^{\mathcal{E}}(\log_2 N)^{3/2}}. \tag{1-2}$$

More generally, consider the restricted multiplication table problem of bounding $M(N; \mathcal{A})$, the number of distinct entries in an $N \times N$ multiplication table that belong to the set \mathcal{A} . For example, when $\lambda \neq 0$ is fixed and $\mathcal{A} = \{p + \lambda : p \text{ prime}\}$, the order of $M(N; \mathcal{A})$ was determined in [5, Theorem 6] (upper bound) and [9] (lower bound).

Observe that $M(N; \mathcal{R}_w) = 1$ when $w \geq N$.

COROLLARY 1.4. *Uniformly for $4 \leq w \leq N/2$ and $\log w \geq (\log N)^{1-(1/\log 4)}$,*

$$M(N; \mathcal{R}_w) \asymp \frac{N^2}{\log^2 w},$$

and when $\log w = (\log y)^\delta$ with $\delta \leq 1 - (1/\log 4)$,

$$M(N; \mathcal{R}_w) \asymp N^2 \delta B(w, N) (\log N)^{-\varepsilon + ((\log(1-\delta))/\log 2)}.$$

PROOF. If $\sqrt{N} < w \leq N/2$, then $M(N; \mathcal{R}_w)$ counts entries in the multiplication table, which are primes in $(w, N]$ or the product of two such primes. The desired bounds follow. If $4 \leq w \leq \sqrt{N}$, we use the inequalities

$$H\left(\frac{N^2}{4}, \frac{N}{4}, \frac{N}{2}; \mathcal{R}_w\right) \leq M(N; \mathcal{R}_w) \leq \sum_{k \geq 0} H\left(\frac{N^2}{2^k}, \frac{N}{2^{k+1}}, \frac{N}{2^k}; \mathcal{R}_w\right).$$

The proof is easy: consider $ab \in \mathcal{R}_w, a \leq N$ and $b \leq N$. If $N/4 < a \leq N/2$ and $ab \leq N^2/4$, then $b \leq N$ and this proves the lower bound. The upper bound comes from taking $N/2^{k+1} < a \leq N/2^k$ for some nonnegative integer k . The desired bound for $M(N; \mathcal{R}_w)$ now follow from Theorem 1.1, since we have $H(x, y, 2y; \mathcal{R}_w) \asymp xf(y, w)$ where $f(u, w) \asymp f(y, w)$ for $\log u \asymp \log y$. □

Next, we consider the ‘Farey fraction multiplication table’. Let \mathcal{F}_N of Farey fractions of order N , that is,

$$\mathcal{F}_N = \left\{ \frac{a}{b} : 1 \leq a \leq b \leq N, (a, b) = 1 \right\}.$$

In private conversation, Igor Shparlinski asked the author about the size of the product set $\mathcal{F}_N \mathcal{F}_N$ (in general, for sets $\mathcal{A}, \mathcal{B} \subseteq \mathbb{Z}$, $\mathcal{A}\mathcal{B}$ denotes the product set $\{ab : a \in \mathcal{A}, b \in \mathcal{B}\}$).

COROLLARY 1.5. *We have*

$$M(N)^2 \ll |\mathcal{F}_N \mathcal{F}_N| \leq M(N)^2.$$

Consequently, by (1-2),

$$|\mathcal{F}_N \mathcal{F}_N| \asymp \frac{N^2}{(\log N)^{\varepsilon} (\log_2 N)^{3/2}}.$$

PROOF. The upper bound is trivial, and thus the real work is on the lower bound. We achieve this by placing restrictions on the fractions, firstly by putting them in dyadic intervals and secondly by removing those elements divisible by small primes. To this end, define

$$\mathcal{A}_N = \{n : N/2 \leq n \leq N\}, \quad \mathcal{A}_N^{(w)} = \mathcal{A}_N \cap \mathcal{R}_w.$$

Let w be a large, fixed constant. A simple inclusion–exclusion argument yields (here p denotes a prime in the sums)

$$\begin{aligned} |\mathcal{F}_N \mathcal{F}_N| &\geq \left| \left\{ \frac{a_1 a_2}{b_1 b_2} : a_1, a_2 \in \mathcal{A}_{N/2}^{(w)}; b_1, b_2 \in \mathcal{A}_N^{(w)}; (a_1 a_2, b_1 b_2) = 1 \right\} \right| \\ &\geq |\mathcal{A}_{N/2}^{(w)} \mathcal{A}_{N/2}^{(w)}| \cdot |\mathcal{A}_N^{(w)} \mathcal{A}_N^{(w)}| - \sum_{w < p \leq N/2} |\mathcal{A}_{N/2}^{(w)} \mathcal{A}_{N/2p}^{(w)}| \cdot |\mathcal{A}_N^{(w)} \mathcal{A}_{N/p}^{(w)}| \\ &\geq |\mathcal{A}_{N/2}^{(w)} \mathcal{A}_{N/2}^{(w)}| \cdot |\mathcal{A}_N^{(w)} \mathcal{A}_N^{(w)}| - \sum_{w < p \leq N/2} |\mathcal{A}_{N/2} \mathcal{A}_{N/2p}| \cdot |\mathcal{A}_N \mathcal{A}_{N/p}|. \end{aligned}$$

It is clear that for $M \leq N$,

$$|\mathcal{A}_N \mathcal{A}_M| \leq H(MN, M/2, M)$$

and we deduce from (1-1) that

$$\begin{aligned} \sum_{w < p \leq N/2} |\mathcal{A}_{N/2} \mathcal{A}_{N/2p}| \cdot |\mathcal{A}_N \mathcal{A}_{N/p}| &\ll \sum_{p > w} \frac{N^4}{p^2 (\log(N/p))^{2\mathcal{E}} (\log_2(N/p))^3} \\ &\ll \frac{M(N)^2}{w \log w}. \end{aligned}$$

We also have the lower bound

$$|\mathcal{A}_N^{(w)} \mathcal{A}_M^{(w)}| \geq H(MN, M/2, M; \mathcal{R}_w) - H(MN/2, M/2, M; \mathcal{R}_w).$$

It follows that

$$\begin{aligned} |\mathcal{F}_N \mathcal{F}_N| &\geq \left(H\left(\frac{N^2}{4}, \frac{N}{4}, \frac{N}{2}; \mathcal{R}_w\right) - H\left(\frac{N^2}{8}, \frac{N}{4}, \frac{N}{2}; \mathcal{R}_w\right) \right) \\ &\quad \times \left(H\left(N^2, \frac{N}{2}, N; \mathcal{R}_w\right) - H\left(\frac{N^2}{2}, \frac{N}{2}, N; \mathcal{R}_w\right) \right) \\ &\quad - O\left(\frac{N^4}{(\log N)^{2\mathcal{E}} (\log_2 N)^3 (w \log w)}\right). \end{aligned} \tag{1-3}$$

Inserting Theorem 1.1 into the estimate (1-3), and taking w to be a sufficiently large constant, we obtain the lower bound in Corollary 1.5. □

1.1. Notation. Let $\tau(n)$ be the number of positive divisors of n , and $\tau(n; y, z)$ denotes the number of divisors of n within the interval $(y, z]$. Let $\omega(n)$ be the number of distinct prime divisors of n . Let $P^+(n)$ be the largest prime factor of n and let $P^-(n)$ be the smallest prime factor of n . Adopt the notational conventions $P^+(1) = 0$ and $P^-(1) = \infty$. Constants implied by O , \ll and \asymp are absolute. The notation $f \asymp g$ means $f \ll g$ and $g \ll f$. The symbol p will always denote a prime. Lastly, $\log_2 x$ denotes $\log \log x$.

1.2. Heuristics. Here we give a short heuristic argument to justify the formulas in Theorem 1.1. This is similar to the heuristics given in [5, 6].

Write $n = n'n''$, where n' is composed only of primes in $(w, 2y]$ and n'' is composed only of primes $> 2y$. For simplicity, assume n' is squarefree and $n' \leq y^{100}$. Assume for the moment that the set $D(n') = \{\log d : d \mid n'\}$ is approximately uniformly distributed in $[0, \log n']$. If n' has k prime factors, then $\tau(n') = 2^k$ and we thus expect that $\tau(n', y, 2y) \geq 1$ with probability about

$$\min\left(1, \frac{2^k}{\log y}\right).$$

This expression changes behavior at $k = k_0 := \lfloor \log_2 y / \log 2 \rfloor$. The number of $n \leq x$ with $n' \in \mathcal{R}_w$ and $\omega(n') = k$ is of size

$$\frac{x}{\log y} \frac{(\log_2 y - \log_2 w)^k}{k!},$$

and we obtain a heuristic estimate for $H(x, y, 2y; \mathcal{R}_w)$ of order

$$\frac{x}{\log^2 y} \left[\sum_{k \leq k_0} \frac{(2 \log_2 y - 2 \log_2 w)^k}{k!} + (\log y) \sum_{k \geq k_0} \frac{(\log_2 y - \log_2 w)^k}{k!} \right].$$

The first sum always dominates, since the second sum is dominated by the first summand (k_0 is always much larger than $\log_2 y - \log_2 w$). The behavior of the first sum over k depends on the relative sizes of k_0 and $2 \log_2 y - 2 \log_2 w$. If $k_0 > 2 \log_2 y - 2 \log_2 w$, that is, $\log w \geq (\log y)^{1-1/\log 4}$, the first contains the ‘peak’ and

$$H(x, y, 2y; \mathcal{R}_w) \approx \frac{x}{\log^2 y} e^{2 \log_2 y - 2 \log_2 w} = \frac{x}{\log^2 w}.$$

For smaller w , we are summing the left tail of the Poisson distribution and standard bounds (see, for example, Lemma 2.4 below) yield

$$H(x, y, 2y; \mathcal{R}_w) \approx xB(y, w)(\log y)^{-\mathcal{E} + ((\log(1-\delta))/\log 2)}.$$

This latter expression is too large by a factor $1/\delta$, and this stems from the uniformity assumption about $D(n')$, which turns out to be false for all but a proportion δ of these integers. Fluctuations in the distribution of the prime factors of n' lead to clustering of the divisors; more details can be found in [5, 6]. As in [5, 6], we really should be considering those n' , which have nicely distributed divisors, and a useful measure of how nicely distributed the divisors are is the function

$$L(a) = \text{meas } \mathcal{L}(a), \quad \mathcal{L}(a) = \bigcup_{d \mid a} [-\log 2 + \log d, \log d).$$

Adjusting our heuristic, we see that the probability that $\tau(n', y, 2y) \geq 1$ should be about $L(n')/\log y$, which is $\gg 1/\log y$ on a set of n' of density δ .

2. Preliminaries

LEMMA 2.1 [6, Lemma 3.1].

- (i) $L(a) \leq \min(\tau(a) \log 2, \log 2 + \log a)$;
- (ii) if $(a, b) = 1$, then $L(ab) \leq \tau(b)L(a)$;
- (iii) if $p_1 < \dots < p_k$, then

$$L(p_1 \cdots p_k) \leq \min_{0 \leq j \leq k} 2^{k-j}(\log(p_1 \cdots p_j) + \log 2).$$

Let $\mathcal{P}(a, b)$ be the set of all squarefree positive integers composed only of primes in $(a, b]$. We adopt the convention that $1 \in \mathcal{P}(a, b)$ for any a, b .

LEMMA 2.2.

- (a) For $t \geq w \geq 2$ and $k \geq 0$,

$$\sum_{\substack{a \in \mathcal{P}(w,t) \\ \omega(a)=k}} \frac{1}{a} \leq \frac{(\log_2 t - \log_2 w + O(1))^k}{k!}.$$

- (b) For $t \geq w \geq 2$ and $k \geq 1$,

$$\sum_{\substack{a \in \mathcal{P}(w,t) \\ \omega(a)=k}} \frac{\log a}{a} \ll (1 + \log(t/w)) \frac{(\log_2 t - \log_2 w + O(1))^{k-1}}{(k-1)!}.$$

- (c) For $2 \leq w \leq s \leq t$,

$$\sum_{a \in \mathcal{P}(w,t)} \frac{L(a)}{a} \ll \left(\frac{\log t}{\log s}\right)^2 \sum_{a \in \mathcal{P}(w,s)} \frac{L(a)}{a}.$$

PROOF. Item (a) is immediate from

$$\sum_{\substack{a \in \mathcal{P}(w,t) \\ \omega(a)=k}} \frac{1}{a} \leq \frac{1}{k!} \left(\sum_{w < p \leq t} \frac{1}{p}\right)^k$$

and Mertens' estimates. For item (b),

$$\sum_{\substack{a \in \mathcal{P}(w,t) \\ \omega(a)=k}} \frac{\log a}{a} = \sum_{\substack{a \in \mathcal{P}(w,t) \\ \omega(a)=k}} \frac{1}{a} \sum_{p|a} \log p \leq \sum_{w < p \leq t} \frac{\log p}{p} \sum_{\substack{a \in \mathcal{P}(w,t) \\ \omega(a)=k-1}} \frac{1}{a}.$$

The desired inequality follows from part (a) and Mertens' estimates. For part (c), we factor each $a \in \mathcal{P}(w, t)$ uniquely as $a = a_1 a_2$ with $a_1 \in \mathcal{P}(w, s)$ and $a_2 \in \mathcal{P}(s, t)$. Then, using Lemma 2.1(ii) we deduce that

$$\begin{aligned} \sum_{a \in \mathcal{P}(w,t)} \frac{L(a)}{a} &\leq \sum_{a_1 \in \mathcal{P}(w,s)} \frac{L(a_1)}{a_1} \sum_{a_2 \in \mathcal{P}(s,t)} \frac{\tau(a_2)}{a_2} \\ &= \prod_{s < p \leq t} \left(1 + \frac{2}{p}\right) \sum_{a_1 \in \mathcal{P}(w,s)} \frac{L(a_1)}{a_1}. \end{aligned}$$

The desired inequality follows from Mertens' estimates. □

The following is a standard sieve bound, see, for example, [8].

LEMMA 2.3. (a) *Uniformly for $x \geq 2z \geq 4$,*

$$|\{x/2 < n \leq x : P^-(n) > z\}| \gg \frac{x}{\log z}.$$

Uniformly for $x \geq z \geq 2$,

$$|\{n \leq x : P^-(n) > z\}| \ll \frac{x}{\log z}.$$

Finally, we quote standard bounds on the Poisson distribution, see, for example, the results in [11, Section 4].

LEMMA 2.4. *Uniformly for $h \leq m \leq x$,*

$$\sum_{h \leq k \leq m} \frac{x^k}{k!} \asymp \min\left(\sqrt{x}, \frac{x}{x-m}, m-h+1\right) \frac{x^m}{m!}.$$

3. Local-to-global estimates

Following a kind of local-to-global principle first utilized in [5], we bound $H(x, y, 2y; \mathcal{R}_w)$ in terms of the function $L(a)$. This justifies the heuristic presented in Section 1.2.

LEMMA 3.1. *If $w \leq y^{1/15}$ and $y \leq \sqrt{x}$, then*

$$H(x, y, 2y; \mathcal{R}_w) - H(x/2, y, 2y; \mathcal{R}_w) \gg \frac{x}{\log^2 y} \sum_{a \in \mathcal{P}(w,y)} \frac{L(a)}{a}.$$

If $w \leq y \leq \sqrt{x}$ and $w \leq y^{1/10}$, then

$$H(x, y, 2y; \mathcal{R}_w) \ll \frac{x}{\log^2 y} \sum_{a \in \mathcal{P}(w,y)} \frac{L(a)}{a}.$$

PROOF. We begin with the lower bound. We may assume without loss of generality that $y \geq y_0$, where y_0 is a sufficiently large constant, because in the case $y < y_0$, for any prime $p \in (y, 2y]$ (such p exists by Bertrand’s postulate) and we see that

$$H(x, y, 2y; \mathcal{R}_w) - H(x/2, y, 2y; \mathcal{R}_w) \gg x/p \gg_{y_0} x.$$

Consider integers $n = ap_1 p_2 b \in (x/2, x]$ with $P^-(a) > w$, p_1 and p_2 prime, satisfying the inequalities

$$a \leq y^{1/5} < p_1 < p_2 \leq \frac{1}{4}y^{4/5} < P^-(b),$$

and with $\log(y/p_1 p_2) \in \mathcal{L}(a)$. The last condition implies that

$$\tau(ap_1 p_2, y, 2y) \geq 1,$$

and we also have that $P^-(n) > w$. Since $y^{4/5} \leq y/a < p_1 p_2 \leq 2y$, we have $x/ap_1 p_2 \geq x/(2y^{6/5}) \geq \frac{1}{2}y^{4/5}$. Thus, by Lemma 2.3, for each triple (a, p_1, p_2) , the number of possible b is $\gg x/(ap_1 p_2 \log y)$. Now $\mathcal{L}(a)$ is the disjoint union of intervals of length $\geq \log 2$ contained in $[-\log 2, \log a]$. For each such interval $[u, v)$, Mertens' estimate implies that

$$\sum_{\substack{u \leq \log(y/p_1 p_2) < v \\ y^{1/5} < p_1 < p_2 < \frac{1}{4}y^{4/5}}} \frac{1}{p_1 p_2} \geq \sum_{8y^{1/5} < p_1 < y^{2/5}} \frac{1}{p_1} \sum_{ye^{-v}/p_1 < p_2 \leq ye^{-u}/p_1} \frac{1}{p_2} \gg \frac{v-u}{\log y}.$$

Here we made use of the estimate $v \leq \log a \leq \frac{1}{5} \log y$, which implies that $ye^{-v}/p_1 \geq y^{2/5} > p_1$. Thus, with a fixed, the sum of $1/(p_1 p_2)$ is $\gg (L(a))/\log y$ and

$$H(x, y, 2y; \mathcal{R}_w) - H(x/2, y, 2y; \mathcal{R}_w) \gg \frac{x}{\log^2 y} \sum_{\substack{a \leq y^{1/5} \\ P^-(a) > w}} \frac{L(a)}{a}.$$

We replace the sum over a with an unbounded set, which is multiplicatively more convenient, starting with

$$\sum_{\substack{a \leq y^{1/5} \\ P^-(a) > w}} \frac{L(a)}{a} \geq \sum_{\substack{a \leq y^{1/5} \\ a \in \mathcal{P}(w, y^{1/15})}} \frac{L(a)}{a} \geq \sum_{a \in \mathcal{P}(w, y^{1/15})} \frac{L(a)}{a} \left(1 - \frac{\log a}{\log(y^{1/5})}\right).$$

Break this into two sums, the first being what we want and the second involving

$$\begin{aligned} \sum_{a \in \mathcal{P}(w, y^{1/15})} \frac{L(a) \log a}{a} &= \sum_{a \in \mathcal{P}(w, y^{1/15})} \frac{L(a)}{a} \sum_{p|a} \log p \\ &= \sum_{w < p \leq y^{1/15}} \frac{\log p}{p} \sum_{\substack{b \in \mathcal{P}(w, y^{1/15}) \\ p \nmid b}} \frac{L(pb)}{b}. \end{aligned}$$

Using the trivial relation $L(pb) \leq 2L(b)$, which comes from Lemma 2.1(ii), and Mertens' estimate,

$$\begin{aligned} \sum_{\substack{a \leq y^{1/5} \\ P^-(a) > w}} \frac{L(a)}{a} &\geq \sum_{a \in \mathcal{P}(w, y^{1/15})} \frac{L(a)}{a} \left(1 - \frac{2 \log(y^{1/15}) + O(1)}{\log(y^{1/5})}\right) \\ &\gg \sum_{a \in \mathcal{P}(w, y^{1/15})} \frac{L(a)}{a}. \end{aligned}$$

An application of Lemma 2.2(c) concludes the proof of the lower bound.

For the upper bound, we first relate $H(x, y, 2y; \mathcal{R}_w)$ to $H^*(x, y, 2y; \mathcal{R}_w)$, the number of squarefree integers $n \leq x$ with $P^-(n) > w$ and $\tau(n, y, z) \geq 1$. Write $n = n' n''$, where n' is squarefree, n'' is squarefull and $(n', n'') = 1$. The number of $n \leq x$ with $n'' > \log^{10} y$ is

$$\leq x \sum_{n'' > \log^{10} y} \frac{1}{n''} \ll \frac{x}{\log^5 y}.$$

If $n'' \leq \log^{10} y$, then for some $f \mid n''$, n' has a divisor in $(y/f, 2y/f]$, hence

$$H(x, y, 2y; \mathcal{R}_w) \leq \sum_{\substack{n'' \leq \log^{10} y \\ P^-(n'') > w}} \sum_{f \mid n''} H^*\left(\frac{x}{n''}, \frac{y}{f}, \frac{2y}{f}; \mathcal{R}_w\right) + O\left(\frac{x}{\log^5 y}\right). \tag{3-1}$$

Let w_0 be a sufficiently large absolute constant. It suffices to prove the upper bound for $w \geq w_0$, for the case $w < w_0$ follows from the case $w = w_0$. In the sum,

$$y/f \leq y \leq (x/n'')^{1/2} \log^5 y \leq (x/n'')^{5/9}$$

for large enough w_0 . We will show that for $w_0 \leq y_1 \leq x_1^{5/9}$,

$$H^*(x_1, y_1, 2y_1; \mathcal{R}_w) \ll x_1 \max_{t \geq y_1^{3/4}} \frac{1}{\log^2 t} \sum_{a \in \mathcal{P}(w, t)} \frac{L(a)}{a}. \tag{3-2}$$

It follows from (3-2) and (3-1) that

$$\begin{aligned} H(x, y, 2y; \mathcal{R}_w) &\ll \sum_{\substack{n'' \leq \log^{10} y \\ P^-(n'') > w}} \frac{x}{n''} \sum_{f \mid n''} \max_{t \geq (y/f)^{3/4}} \frac{1}{\log^2 t} \sum_{a \in \mathcal{P}(w, t)} \frac{L(a)}{a} \\ &\ll x \max_{t \geq y^{2/3}} \frac{1}{\log^2 t} \sum_{a \in \mathcal{P}(w, t)} \frac{L(a)}{a} \sum_{\substack{n'' \leq \log^{10} y \\ P^-(n'') > w}} \frac{\tau(n'')}{n''}. \end{aligned}$$

The lemma follows by noting that the inner sum over squarefull n'' is $O(1)$, using the relative estimate in Lemma 2.2(c) with $s = y^{2/3}$, and finally noting that $\mathcal{P}(w, y^{2/3}) \subseteq \mathcal{P}(w, y)$.

It remains to prove (3-2). The right side is $\gg x_1/\log^2 y_1$ since $L(1) = \log 2$, and hence it suffices to count those $n \in (x_1/\log^2 y_1, x_1]$. We will count separately those $n \in (x_1/2^{r+1}, x_1/2^r]$ for some integer r , $0 \leq r \leq 5 \log_2 y_1$. Let \mathcal{A} be the set of squarefree integers $n \in (x_1/2^{r+1}, x_1/2^r]$ with a divisor in $(y_1, 2y_1]$. Put $z_1 = 2y_1$, $y_2 = x_1/(2^{r+2}y_1)$, $z_2 = x_1/(2^r y_1)$. If $n \in \mathcal{A}$, then $n = m_1 m_2$ with $y_i < m_i \leq z_i$ ($i = 1, 2$). For some $j \in \{1, 2\}$ we have $p = P^+(m_j) < P^+(m_{3-j})$; in particular, p is not the largest prime factor of n . Fixing j , we may write $n = abp$, where $P^+(a) < p < P^-(b)$ and $b > p$. Since $\tau(ap, y_j, z_j) \geq 1$, we have $y_j/a \leq p \leq z_j$. By Lemma 2.3 and the fact that $b > p$, given a and p , the number of choices for b is

$$\begin{aligned} &\ll \frac{x_1}{2^r a p \log p} \\ &\leq \frac{x_1}{2^r a p \log \max(P^+(a), y_j/a)}. \end{aligned}$$

Now a has a divisor in $(y_j/p, z_j/p]$, and thus

$$\log(y_j/p) \in \mathcal{L}(a) \quad \text{or} \quad \log(2y_j/p) \in \mathcal{L}(a).$$

Since $\mathcal{L}(a)$ is the disjoint union of intervals of length $\geq \log 2$ with total measure $L(a)$, by repeated use of Mertens' estimate

$$\sum_{\substack{\log(cy_j/p) \in \mathcal{L}(a) \\ p \geq P^+(a)}} \frac{1}{p} \ll \frac{L(a)}{\log \max(P^+(a), y_j/a)} \quad (c = 1, 2).$$

Since $y_j \geq y_1^{4/9} / 2^{r+2} \geq y_1^{3/4}$,

$$H^*(x, y, 2y; \mathcal{R}_w) \ll \sum_{0 \leq r \leq 5 \log_2 y_1} \frac{x_1}{2^r} \sum_{t \in \{4y_1, 4y_2\}} \sum_{a \in \mathcal{P}(w,t)} \frac{L(a)}{a \log^2(P^+(a) + \frac{t}{4a})}.$$

We have $4y_j \geq y_1^{4/5} / 2^r \geq y_1^{3/4}$ for any j and any r . Also, by [10, Lemma 2.2],

$$\sum_{a \in \mathcal{P}(w,t)} \frac{L(a)}{a \log^2(t/(4a) + P^+(a))} \ll \frac{1}{\log^2 t} \sum_{a \in \mathcal{P}(w,t)} \frac{L(a)}{a}.$$

Summing over r , we deduce (3-2). □

4. Proof of Theorem 1.1: lower bounds

We first deal with simple cases. Let w_0 be a sufficiently large constant and $\varepsilon > 0$ a sufficiently small constant. Firstly, if $y \leq w_0$, then Bertrand's postulate implies that there is a prime $p \in (y, 2y]$ and therefore

$$H(x, y, 2y; \mathcal{R}_w) - H_z(x/2, y, 2y; \mathcal{R}_w) \geq \#\{x/2 < n \leq x : p \mid n\} \gg x.$$

Also, if $w \leq w_0 < y$ and $w \leq y/8$, then

$$H(x, y, 2y; \mathcal{R}_w) - H(x/2, y, 2y; \mathcal{R}_w) \geq H(x, y, 2y; \mathcal{R}_{w_0}) - H(x/2, y, 2y; \mathcal{R}_{w_0})$$

and the desired bound follows from the case $w = w_0$. Thirdly, when $y > w_0$ and $y^\varepsilon < w \leq y/8$, we consider two cases: (a) $y \leq \sqrt{x/8}$ and (b) $\sqrt{x/8} < y \leq \sqrt{x}$. In case (a), consider $n = pm$ where $y < p \leq 2y < P^-(m)$. Since $x/p \geq 4y$ for all such p , Lemma 2.3 implies

$$\begin{aligned} H(x, y, 2y; \mathcal{R}_w) - H(x/2, y, 2y; \mathcal{R}_w) &\geq \sum_{y < p \leq 2y} \#\left\{ \frac{x}{2p} < n \leq \frac{x}{p} : P^-(n) > 2y \right\} \\ &\gg \sum_{y < p \leq 2y} \frac{x}{p \log y} \gg \frac{x}{\log^2 w}. \end{aligned}$$

In the case (b) $\sqrt{x/8} < y \leq \sqrt{x}$, consider $n = pm$ where $y < p \leq 2y$ and $P^-(m) > y/8$. Such n have at most three prime factors larger than y , hence at most three representations in this form. Since $x/p \geq 2y/8$, Lemma 2.3 similarly implies that

$$\begin{aligned} H(x, y, 2y; \mathcal{R}_w) - H(x/2, y, 2y; \mathcal{R}_w) &\geq \frac{1}{3} \sum_{y < p \leq 2y} \#\left\{ \frac{x}{2p} < n \leq \frac{x}{p} : P^-(n) > \frac{y}{8} \right\} \\ &\gg \sum_{y < p \leq 2y} \frac{x}{p \log y} \gg \frac{x}{\log^2 w}. \end{aligned}$$

From now on, we assume

$$w_0 < w \leq y^\varepsilon. \tag{4-1}$$

We begin with the local-to-global estimate for $H(x, y, 2y; \mathcal{R}_w)$ given in Lemma 3.1, and relate $L(a)$ to counts of *pairs* of divisors, which are close together. Evidently,

$$L(a) \geq (\log 2) \#\{d \mid a : \tau(a, d, 2d) = 0\} \geq (\log 2)(\tau(a) - W^*(a)), \tag{4-2}$$

where

$$W^*(a) = \#\{d \mid a, d' \mid a : d < d' \leq 2d\}.$$

We will apply (4-2) with integers whose prime factors are localized. As in [6], partition the primes into sets D_1, D_2, \dots , where each D_j consists of the primes in an interval $(\lambda_{j-1}, \lambda_j]$, with $\lambda_j \approx \lambda_{j-1}^2$. More precisely, let $\lambda_0 = 1.9$ and define inductively λ_j for $j \geq 1$ as the largest prime so that

$$\sum_{\lambda_{j-1} < p \leq \lambda_j} \frac{1}{p} \leq \log 2. \tag{4-3}$$

For example, $\lambda_1 = 2$ and $\lambda_2 = 7$. By Mertens' bounds,

$$\log_2 \lambda_j - \log_2 \lambda_{j-1} = \log 2 + O(1/\log \lambda_{j-1}),$$

and it follows that for some absolute constant K ,

$$2^{j-K} \leq \log \lambda_j \leq 2^{j+K} \quad (j \geq 0). \tag{4-4}$$

For a vector $\mathbf{b} = (b_1, \dots, b_j)$ of nonnegative integers, let $\mathcal{A}(\mathbf{b})$ be the set of squarefree integers a composed of exactly b_j prime factors from D_j for each j .

LEMMA 4.1. *Assume $\mathbf{b} = (b_1, \dots, b_{J_2})$, with $b_j = 0$ for $j < J_1$. Then*

$$\sum_{a \in \mathcal{A}(\mathbf{b})} \frac{W^*(a)}{a} \ll \frac{(2 \log 2)^{b_{J_1} + \dots + b_{J_2}}}{b_{J_1}! \cdots b_{J_2}!} \sum_{j=J_1}^{J_2} 2^{-j+b_{J_1} + \dots + b_j}.$$

PROOF. Identical to the proof of [6, Lemma 2.3], except that we remove the terms corresponding to $d = d'$. □

We will only consider those intervals $D_j \subseteq (w, y]$, that is, only $J_1 \leq j \leq J_2$, where

$$J_1 := \min\{j : \lambda_{j-1} > w\}, \quad J_2 := \max\{j : \lambda_j \leq y\}. \tag{4-5}$$

By (4-4),

$$\left| J_1 - \frac{\log_2 w}{\log 2} \right| \leq K + 2, \quad \left| J_2 - \frac{\log_2 y}{\log 2} \right| \leq K + 1. \tag{4-6}$$

Put

$$M = \frac{\log_2 w_0}{200}, \tag{4-7}$$

which is a sufficiently large constant because w_0 is. Recalling (4-1) and (4-5), we may choose ε such that

$$J_1 \geq 100M, \quad J_2 - J_1 \geq 100M. \tag{4-8}$$

Let \mathcal{B}_k be the set of vectors $(b_{J_1}, \dots, b_{J_2})$ which satisfy the following:

- (a) $b_{J_1} + \dots + b_{J_2} = k$;
- (b) $\sum_{j=J_1}^{J_2} 2^{-j+b_{J_1}+\dots+b_j} \leq 2^{-M}$;
- (c) $b_{J_1+i-1} \leq M + i^2$ ($i \geq 1$);
- (d) $b_{J_2-i+1} \leq M + i^2$ ($i \geq 1$).

Item (b) ensures that the sum on a in Lemma 4.1 is small, provided that w_0 is sufficiently large. From the definition of J_2 , whenever $\mathbf{b} \in \mathcal{B}_k$ and $a \in \mathcal{A}(\mathbf{b})$, we have $a \in \mathcal{P}(w, y)$.

By Lemma 4.1, for any k and any $\mathbf{b} \in \mathcal{B}_k$,

$$\sum_{a \in \mathcal{A}(\mathbf{b})} \frac{W^*(a)}{a} \leq \frac{1}{10} \frac{(\log 4)^k}{b_{J_1}! \dots b_{J_2}!}. \tag{4-9}$$

By (4-4), the fact that J_1 is sufficiently large, and $b_j \leq (j + 1 - J_1)^2 + M$, for any k and $\mathbf{b} \in \mathcal{B}_k$ we have by (4-3)

$$\begin{aligned} \sum_{a \in \mathcal{A}(\mathbf{b})} \frac{\tau(a)}{a} &= 2^k \prod_{j=J_1}^{J_2} \frac{1}{b_j!} \left(\sum_{p_1 \in D_j} \frac{1}{p_1} \sum_{\substack{p_2 \in D_j \\ p_2 \neq p_1}} \frac{1}{p_2} \dots \sum_{\substack{p_{b_j} \in D_j \\ p_{b_j} \notin \{p_1, \dots, p_{b_j-1}\}}} \frac{1}{p_{b_j}} \right) \\ &\geq 2^k \prod_{j=J_1}^{J_2} \frac{1}{b_j!} \left(\log 2 - \frac{b_j}{\lambda_{j-1}} \right)^{b_j} \geq \frac{(\log 4)^k}{2b_{J_1}! \dots b_{J_2}!}. \end{aligned} \tag{4-10}$$

Combining Lemma 3.1, (4-2), (4-9), and (4-10), we arrive at

$$H(x, y, 2y; \mathcal{R}_w) - H(x/2, y, 2y; \mathcal{R}_w) \geq \frac{x}{\log^2 y} \sum_{k_1 \leq k \leq k_2} \sum_{\mathbf{b} \in \mathcal{B}_k} \frac{1}{b_{J_1}! \dots b_{J_2}!} \tag{4-11}$$

for any $k_1 \leq k_2$. We bound the sum on \mathbf{b} using techniques from [5].

Following our heuristic, take

$$k_2 = \left\lfloor \min \left(\frac{\log_2 y}{\log 2}, 2(\log_2 y - \log_2 w) \right) - 2M \right\rfloor. \tag{4-12}$$

By (4-1), $k_2 \geq 100M$ and by (4-6),

$$k_2 = \min(J_2, (\log 4)(J_2 - J_1)) - 2M + \theta, \quad |\theta| \leq (\log 4)(2K + 3).$$

We will choose k_1 to satisfy

$$10M \leq k_1 \leq k_2. \tag{4-13}$$

Also define

$$v = J_2 - J_1 + 1, \quad s = J_1 - 2 - M. \tag{4-14}$$

Setting $g_i = b_{J_1+i-1}$ for $i \geq 1$,

$$\sum_{i=1}^v 2^{-i+g_1+\dots+g_i} = 2^{J_1-1} f(\mathbf{b}) \leq 2^{s+1}.$$

By (c) and (d) in the definition of \mathcal{B}_k , $g_i \leq M + i^2$ and $g_{v+1-i} \leq M + i^2$ for every $i \geq 1$. Applying the argument on the top of [5, page 419], it follows that for $k_1 \leq k \leq k_2$

$$\sum_{\mathbf{b} \in \mathcal{B}_k} \frac{(\log 4)^k}{b_{J_1}! \cdots b_{J_2}!} \gg v^k \text{Vol}(Y_k(s, v)), \tag{4-15}$$

where $Y_k(s, v)$ is the set of $\xi = (\xi_1, \dots, \xi_k) \in \mathbb{R}^k$ satisfying the following:

- (i) $0 \leq \xi_1 \leq \dots \leq \xi_k < 1$;
- (ii) For $1 \leq i \leq \sqrt{k - M}$, $\xi_{M+i^2} > i/v$ and $\xi_{k+1-(M+i^2)} < 1 - i/v$;
- (iii) $\sum_{j=1}^k 2^{j-v\xi_j} \leq 2^s$.

We now invoke a result from [5] concerning the volume of $Y_k(s, v)$.

LEMMA 4.2 [5, Lemma 4.9]. *The following is true for all sufficiently large M . Uniformly for $v \geq 1$, $10M \leq k \leq 100(v - 1)$, $s \geq M/2 + 1$ and $0 \leq k - v \leq s - M/3 - 1$. Then*

$$\text{Vol}(Y_k(s, v)) \gg \frac{k - v + 1}{(k + 1)!}.$$

If w_0 is large enough (implying that M is sufficiently large) and ε is sufficiently small, then (4-5), (4-7), (4-12), (4-13), and (4-14) together imply that

$$\begin{aligned} v &= J_2 - J_1 + 1 \geq 10, \\ 10M &\leq k_1 \leq k_2 \leq (\log 4)(J_2 - J_1) = (\log 4)(v - 1), \\ s &\geq \log_2 w - M \geq M/2 + 1, \\ k_2 - v - s &\leq (J_2 - 2M) - (J_2 - 1 - M) = 1 - M \leq -M/3 - 1. \end{aligned}$$

Thus, we see that the hypotheses of Lemma 4.2 are satisfied. Therefore, gathering (4-11), (4-15) and invoking Lemma 4.2, we conclude that

$$H(x, y, 2y; \mathcal{R}_w) - H(x/2, y, 2y; \mathcal{R}_w) \gg \frac{x}{\log^2 y} \sum_{k=k_1}^{k_2} \frac{(v \log 4)^k}{k!} \left(\frac{k - v + 1}{k + 1} \right). \tag{4-16}$$

Consider three cases: I. $\delta \geq 1 - 1/\log 4$, II. $\frac{1}{10} \leq \delta < 1 - 1/\log 4$, III. $0 < \delta < \frac{1}{10}$.

Case I. We have $\log_2 w \geq (1 - 1/\log 4)\log_2 y$ and thus, by (4-12) and (4-5),

$$k_2 = \lfloor 2(\log_2 y - \log_2 w) - 2M \rfloor.$$

For an appropriate choice of parameters ε, M, w_0 (subject to (4-7)), we have by (4-8) the bound

$$v \log 4 \geq k \geq J_2 - J_1 \geq 100M.$$

Now set $k_1 = k_2/10$, so that (4-13) is satisfied. With these choices and (4-14),

$$\frac{k - v + 1}{k + 1} \asymp 1 \quad (k_1 \leq k \leq k_2).$$

Thus, apply Lemma 2.4 to the sum in (4-16),

$$H(x, y, 2y; \mathcal{R}_w) - H(x/2, y, 2y; \mathcal{R}_w) \gg \frac{x}{\log^2 y} e^{v \log 4} \gg \frac{x}{\log^2 w},$$

as required in this case.

Case II. By (4-12),

$$k_2 = \left\lfloor \frac{\log_2 y}{\log 2} - 2M \right\rfloor,$$

and take

$$k_1 = \frac{k_2}{2}.$$

In this case, we have $J_1 \asymp \log_2 y \asymp J_2 - J_1$, and thus by (4-14),

$$\frac{k - v + 1}{k + 1} \asymp 1 \asymp \delta \quad (k_1 \leq k \leq k_2).$$

Hence, applying Lemma 2.4 with $h = k_1, m = k_2, x = v \log 4$, we compute

$$\begin{aligned} \min\left(x^{1/2}, \frac{x}{x - m}, m - h + 1\right) &\gg \min\left((\log_2 y)^{1/2}, \frac{v \log 4}{v \log 4 - k_2}\right) \\ &\gg \min\left((\log_2 y)^{1/2}, \frac{1}{1 - \frac{1}{\log 4} - \delta} + O(1/\log_2 y)\right) \\ &\gg \delta(\log_2 y)^{1/2} B(w, y). \end{aligned}$$

Recalling the definition of \mathcal{E} , by Stirling's formula,

$$\frac{(v \log 4)^{k_2}}{k_2!} \gg \frac{(e(1 - \delta))^{k_2}}{\sqrt{\log_2 y}} = \frac{(\log y)^{2 - \mathcal{E} + (\log(1 - \delta))/(\log 2)}}{\sqrt{\log_2 y}}.$$

Invoking Lemma 2.4 we see that the sum in (4-16) is

$$\gg \delta B(w, y) (\log y)^{2 - \mathcal{E} + (\log(1 - \delta))/(\log 2)},$$

and this gives the required lower bound in Theorem 1.1.

Case III. When $\delta < \frac{1}{10}$, also

$$k_2 = \left\lfloor \frac{\log_2 y}{\log 2} - 2M \right\rfloor,$$

but in this case,

$$k_1 = k_2,$$

as we are in the range where the summation in (4-16) is dominated by single summand.

Here

$$k_2 - v + 1 \asymp J_1, \quad \frac{k - v + 1}{k + 1} \asymp \frac{\log_2 w}{\log_2 y} = \delta.$$

Applying Lemma 2.4 to the sum in (4-16),

$$H(x, y, 2y; \mathcal{R}_w) - H(x/2, y, 2y; \mathcal{R}_w) \gg \frac{\delta x}{\log^2 y} \frac{(v \log 4)^{k_2}}{k_2!}.$$

Applying Stirling’s formula as in Case II and observing that $B(w, y) = 1$ in this case, we conclude the desired upper bound.

This completes the proof of the lower bound in Theorem 1.1.

5. Proof of Theorem 1.1: upper bounds

In this section, we prove the upper bound in Theorem 1.1. We begin with simple cases. If w_0 is fixed and $w \leq w_0$, then $H(x, y, 2y; \mathcal{R}_w) \leq H(x, y, 2y)$ and the required bound follows from (1-1). Next, if $\log_2 w \geq (1 - 1/\log 4) \log_2 y$, then by Lemma 2.3,

$$\begin{aligned} H(x, y, 2y; \mathcal{R}_w) &\leq \sum_{\substack{y < d \leq 2y \\ P^-(d) > w}} |\{m \leq x/d : P^-(m) > w\}| \\ &\ll \sum_{\substack{y < d \leq 2y \\ P^-(d) > w}} \frac{x}{d \log w} \\ &\ll \frac{x}{\log^2 w}, \end{aligned}$$

as required.

From now on, we assume that

$$\log w \leq (\log y)^{1-1/\log 4}, \tag{5-1}$$

that is, $\delta \leq 1 - 1/\log 4$. We apply Lemma 3.1 and use upper bounds for $L(a)$ from Lemma 2.1. As in [5], the sums involving $L(a)$ are bounded in terms of multivariate integrals, which were estimated accurately in [5, 6].

5.1. Case I. $\frac{1}{10} \leq \delta \leq 1 - 1/\log 4$. This case is very easy, as we expect no clustering of divisors. Let

$$k_0 = \left\lfloor \frac{\log_2 y}{\log 2} \right\rfloor. \tag{5-2}$$

Beginning with Lemma 3.1, we apply Lemma 2.1(i) to bound $L(a)$ and then apply Lemma 2.2 parts (a) and (b)

$$\begin{aligned} H(x, y, 2y; \mathcal{R}_w) &\ll \frac{x}{\log^2 y} \left[\sum_{k \leq k_0} 2^k \sum_{\substack{a \in \mathcal{P}(w, y) \\ \omega(a) = k}} \frac{1}{a} + \sum_{k > k_0} \sum_{\substack{a \in \mathcal{P}(w, y) \\ \omega(a) = k}} \frac{\log a}{a} \right] \\ &\ll \frac{x}{\log^2 y} \left[\sum_{k \leq k_0} \frac{(2 \log_2 y - 2 \log_2 w)^k}{k!} + (\log y) \sum_{k \geq k_0} \frac{(\log_2 y - \log_2 w)^k}{k!} \right]. \end{aligned}$$

Since $k_0 \geq 1.4(\log_2 y - \log_2 w)$, the second sum on the right side is dominated by the single term $k = k_0$ and thus, by Stirling’s formula,

$$\begin{aligned} \sum_{k \geq k_0} \frac{(\log_2 y - \log_2 w)^k}{k!} &\ll \frac{(\log_2 y - \log_2 w)^{k_0}}{k_0!} \\ &\ll \frac{((e \log 2)(1 - \delta))^{k_0}}{(\log_2 y)^{1/2}} \\ &\ll \frac{(\log y)^{1 - \varepsilon + (\log(1 - \delta))/\log 2}}{(\log_2 y)^{1/2}}. \end{aligned}$$

We have $k_0 \leq 2 \log_2 y - 2 \log_2 w$ in the first sum, for which we invoke Lemma 2.4 and obtain, with $\alpha = \log_2 y - \log_2 w$ the bound

$$\begin{aligned} \sum_{k \leq k_0} \frac{(2 \log_2 y - 2 \log_2 w)^k}{k!} &\ll \frac{(2\alpha)^{k_0}}{k_0!} \min\left(\alpha^{1/2}, \frac{\alpha}{\alpha - k_0}\right) \\ &\ll (2e(\log 2)(1 - \delta))^{k_0} \min(1, (\log_2 y)^{-1/2}((1 - \delta) \log 4 - 1)^{-1}) \\ &\ll (\log y)^{-\varepsilon + (\log(1 - \delta))/\log 2} B(w, y), \end{aligned}$$

as required for Theorem 1.1.

5.2. Case II. $\delta \leq \frac{1}{10}$. This case is more delicate, because we expect that typically there will be clustering of the divisors of a , and we must bound the probability of nonclustering.

We cut up the sum in Lemma 3.1 according to $\omega(a)$. Let

$$T_k = \sum_{\substack{a \in \mathcal{P}(w, y) \\ \omega(a) = k}} \frac{L(a)}{a}.$$

We bound T_k in terms of a multivariate integral, in a manner similar to that in [6].

LEMMA 5.1. *Suppose w is large, (5-1) holds, let*

$$v = \left\lfloor \frac{\log_2 y - \log_2 w}{\log 2} \right\rfloor, \quad u = \left\lfloor \frac{\log_2 w}{\log 2} \right\rfloor$$

and assume that $1 \leq k \leq 10v$. Then

$$T_k \ll (2 \log_2 y - 2 \log_2 w)^k U_k(v, u),$$

where

$$U_k(v, u) = \int \cdots \int_{0 \leq \xi_1 \leq \cdots \leq \xi_k \leq 1} \min_{0 \leq j \leq k} 2^{-j} (2^{v\xi_1 + u} + \cdots + 2^{v\xi_j + u} + 1) d\xi.$$

PROOF. The proof is the same as the [6, proof of Lemma 3.5], except that we make use of the fact that $P^-(a) > w$. Recall the definition of the sets D_j from Section 4. By (4-4), any prime divisor of a lies in D_j with $u - K - 2 \leq j \leq v + u + K + 3$. Following the proof of [6, Lemma 3.5], in particular using Lemma 2.1(iii),

$$T_k \ll \frac{(2 \log 2)^k}{k!} \int_{[u-K-2, v+u+K+4]^k} F(\mathbf{t}) \, d\mathbf{t}, \tag{5-3}$$

where, letting $s_1 \leq s_2 \leq \dots \leq s_k$ be the increasing rearrangement of t_1, \dots, t_k ,

$$F(\mathbf{t}) = \min_{0 \leq j \leq k} 2^{-j} (2^{s_1} + \dots + 2^{s_j} + 1).$$

Observe that $F(\mathbf{t})$ is symmetric in t_1, \dots, t_k . Making the change of variables

$$t_i = u - K - 2 + (v + 2K + 6)\xi_i \quad (1 \leq i \leq k)$$

we see that $0 \leq \xi_i \leq 1$ for each i . Utilizing the symmetry of $F(\mathbf{t})$, we find that the multiple integral on the right side of (5-3) equals

$$(v + 2K + 6)^k k! \int \dots \int_{0 \leq \xi_1 \leq \dots \leq \xi_k \leq 1} \min_{0 \leq j \leq k} 2^{-j} (2^{(v+2K+6)\xi_1+u} + \dots + 2^{(v+2K+6)\xi_j+u} + 1) \, d\xi.$$

We conclude that

$$T_k(y) \ll ((2 \log 2)(v + 2K + 6))^k U_k(v, u).$$

Lastly, $(v + 2K + 6)^k \ll v^k$ since $k \leq 10v$, and the lemma follows. □

To bound $U_k(u, v)$ we invoke the following estimate from [5, 6].

LEMMA 5.2 ([5, Lemma 13.2], [6, Lemma 4.4]). *Define*

$$\mathcal{T}(k, v, \gamma) = \{ \boldsymbol{\xi} \in \mathbb{R}^k : 0 \leq \xi_1 \leq \dots \leq \xi_k \leq 1, \\ 2^{v\xi_1} + \dots + 2^{v\xi_j} \geq 2^{j-\gamma} \quad (1 \leq j \leq k) \}.$$

Suppose $k, v, \gamma \in \mathbb{Z}$ with $1 \leq k \leq 10v$ and $\gamma \geq 0$. Set $b = k - v$. Then

$$\text{Vol}(\mathcal{T}(k, v, \gamma)) \ll \frac{Y}{2^{2b-\gamma} (k + 1)!},$$

where

$$Y = \begin{cases} b & \text{if } b \geq \gamma + 5, \\ (\gamma + 5 - b)^2 (\gamma + 1) & \text{if } b < \gamma + 5. \end{cases}$$

LEMMA 5.3. *Suppose k, u, v are integers satisfying $1 \leq k \leq 10v$ and $u \geq 1$. Then*

$$U_k(v, u) \ll \frac{u(1 + |k - v - u|^2)}{(k + 1)! (2^{k-v-u} + 1)}.$$

Notice that the bound in Lemma 5.3 undergoes a change of behavior at $k = v + u$.

PROOF. Put $b = k - v$. For integers $m \geq 0$, consider ξ satisfying

$$2^{-m} \leq \min_{0 \leq j \leq k} 2^{-j}(2^{v\xi_1+u} + \dots + 2^{v\xi_j+u} + 1) < 2^{1-m}.$$

For $1 \leq j \leq k$,

$$2^{-j}(2^{v\xi_1+u} + \dots + 2^{v\xi_j+u}) \geq \max(2^{-j}, 2^{-m-u} - 2^{-j-u}) \geq 2^{-m-u-1},$$

and thus $\xi \in \mathcal{T}(k, v, m + u + 1)$. Invoking Lemma 5.2,

$$\begin{aligned} U_k(v, u) &\leq \sum_{m \geq 0} 2^{1-m} \text{Vol}(\mathcal{T}(k, v, m + u + 1)) \\ &\ll \frac{1}{(k + 1)!} \sum_{m \geq 0} \frac{2^{-m} Y_m}{2^{2^{b-m-u-1}}}, \\ Y_m &= \begin{cases} b & \text{if } m + u \leq b - 6, \\ (m + u + 6 - b)^2(m + u + 2) & \text{if } m + u > b - 6. \end{cases} \end{aligned}$$

Dividing the sum according to the two cases yields

$$\begin{aligned} \sum_{m \geq 0} \frac{2^{-m} Y_m}{2^{2^{b-m-u-1}}} &\ll \sum_{0 \leq m < b-u-5} \frac{b}{2^m 2^{2^{b-m-u-1}}} \\ &\quad + \sum_{m \geq \max(0, b-u-5)} \frac{(m + u + 6 - b)^2(m + u + 2)}{2^m}. \end{aligned}$$

The proof is completed by noting that if $b \geq 6 + u$, each sum on the right side is $\ll b2^{u-b}$ and if $b \leq 5 + u$, the first sum is empty and the second is $\ll (6 + u - b)^2 \ll 1 + (b - u)^2$. □

Finally, we complete the upper bound in Theorem 1.1. Let

$$v = \left\lfloor \frac{\log_2 y - \log_2 w}{\log 2} \right\rfloor, \quad u = \left\lfloor \frac{\log_2 w}{\log 2} \right\rfloor$$

and define k_0 by (5-2). Note that $k_0 = v + u + O(1)$. We now combine Lemmas 5.1 and 5.3. Since $k_0 > 1.4(\log_2 y - \log_2 w)$,

$$\begin{aligned} \sum_{k_0 \leq k \leq 10k_0} T_k &\ll \sum_{k_0 \leq k \leq 10k_0} \frac{u(1 + (k - k_0)^2)}{(k + 1)! 2^{k-u-v}} (2 \log_2 y - 2 \log_2 w)^k \\ &\ll u 2^{k_0} \sum_{\ell \geq 0} \frac{1 + \ell^2}{(k_0 + 1 + \ell)!} (\log_2 y - \log_2 w)^{k_0 + \ell} \\ &\ll (\log_2 w) \frac{(2 \log_2 y - 2 \log_2 w)^{k_0}}{(k_0 + 1)!}. \end{aligned}$$

Similarly, since $k_0 \leq 0.9(2 \log_2 y - 2 \log_2 w)$,

$$\begin{aligned} \sum_{0 \leq k < k_0} T_k &\ll 1 + \sum_{1 \leq k < k_0} \frac{u(k_0 - k)^2 (2 \log_2 y - 2 \log_2 w)^k}{(k + 1)!} \\ &\ll 1 + u \sum_{\ell=1}^{k_0-1} \frac{u \ell^2 (2 \log_2 y - 2 \log_2 w)^{k_0-\ell}}{(k_0 + 1 - \ell)!} \\ &\ll (\log_2 w) \frac{(2 \log_2 y - 2 \log_2 w)^{k_0}}{(k_0 + 1)!}. \end{aligned}$$

For the large values of k we use the crude bound $L(a) \ll \tau(a)$ from Lemma 2.1(i), followed by an application of Lemma 2.2(a). This gives

$$\begin{aligned} \sum_{k \geq 10k_0} T_k &\leq \sum_{k \geq 10k_0} \sum_{\substack{a \in \mathcal{P}(w,y) \\ \omega(a)=k}} \frac{2^k \log 2}{a} \leq \sum_{k \geq 10k_0} \frac{(2 \log_2 y - 2 \log_2 w + O(1))^k}{k!} \\ &\ll \frac{(2 \log_2 y - 2 \log_2 w + O(1))^{10k_0}}{(10k_0)!} \\ &\ll \frac{(2 \log_2 y - 2 \log_2 w)^{k_0}}{(k_0 + 1)!}. \end{aligned}$$

Combining these three bounds for sums of T_k with Lemma 3.1, Lemma 2.4, and Stirling’s formula, we conclude that

$$\begin{aligned} H(x, y, 2y; \mathcal{R}_w) &\ll \frac{x}{\log^2 y} (\log_2 w) \frac{(2 \log_2 y - 2 \log_2 w)^{k_0}}{(k_0 + 1)!} \\ &\ll \frac{x \log_2 w}{(\log_2 y)^{3/2}} (\log y)^{-\varepsilon + ((\log(1-\delta))/\log 2)}. \end{aligned}$$

The proof of the upper bound in Theorem 1.1 is complete.

References

- [1] P. Erdős, ‘On the greatest prime factor of $\prod_{k=1}^x f(k)$ ’, *J. Lond. Math. Soc. (2)* **27** (1952), 379–384.
- [2] P. Erdős, ‘Some remarks on number theory’, *Riveon Lematematika* **9** (1955), 45–48 (Hebrew. English summary).
- [3] P. Erdős, ‘An asymptotic inequality in the theory of numbers’, *Vestnik Leningrad. Univ. (13)* **15** (1960), 41–49 (in Russian).
- [4] P. Erdős and A. Schinzel, ‘On the greatest prime factor of $\prod_{k=1}^x f(k)$ ’, *Acta Arith. (2)* **55** (1990), 191–200.
- [5] K. Ford, ‘The distribution of integers with a divisor in a given interval’, *Ann. Math.* **168** (2008), 367–433.
- [6] K. Ford, ‘Integers with a divisor in $(y, 2y]$ ’, in: *Anatomy of Integers*, CRM Proceedings and Lecture Notes, 46 (eds. J.-M. De Koninck, A. Granville and F. Luca) (Providence, RI, 2008), 65–80.
- [7] K. Ford, M. R. Khan, I. E. Shparlinski and C. L. Yankov, ‘On the maximal difference between an element and its inverse in residue rings’, *Proc. Amer. Math. Soc. (12)* **133** (2005), 3463–3468.
- [8] H. Halberstam and H.-E. Richert, *Sieve Methods*, London Mathematical Society Monographs, 4 (Academic Press, 1974).

- [9] D. Koukoulopoulos, 'Divisors of shifted primes', *Int. Math. Res. Not. IMRN (24)* **2010** (2010), 4585–4627.
- [10] D. Koukoulopoulos, 'Localized factorizations of integers', *Proc. Lond. Math. Soc.* **101** (2010), 392–426.
- [11] K. K. Norton, 'On the number of restricted prime factors of an integer. I', *Illinois J. Math. (4)* **20** (1976), 681–705.
- [12] G. Tenenbaum, 'Sur une question d'Erdős et Schinzel', in: *A Tribute to Paul Erdős* (Cambridge University Press, Cambridge, 1990), 405–443.
- [13] G. Tenenbaum, 'Sur une question d'Erdős et Schinzel. II', *Invent. Math. (1)* **99** (1990), 215–224.

KEVIN FORD, Department of Mathematics, 1409 West Green Street,
University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA
e-mail: ford126@illinois.edu