INTEGERS WITH A DIVISOR IN $(y, 2y]$

KEVIN FORD

ABSTRACT. We determine, up to multiplicative constants, how many integers $n \leq x$ have a divisor in $(y, 2y]$.

1. INTRODUCTION

Let $H(x, y, z)$ be the number of integers $n \leq x$ which have a divisor in the interval $(y, z]$. In the author's paper [4], the correct order of growth of $H(x, y, z)$ was determined for all x, y, z. In particular,

$$(1.1) \qquad H(x, y, 2y) \asymp \frac{x}{(\log y)^\delta (\log \log y)^{3/2}} \qquad (3 \leq y \leq \sqrt{x}),$$

where

$$\delta = 1 - \frac{1 + \log \log 2}{\log 2} = 0.086071\ldots.$$

In this note we prove only the important special case (1.1), omitting the parts of the argument required for other cases. In addition, we present an alternate proof of the lower bound implicit in (1.1). This proof avoids the use of results about uniform order statistics required in [4], and instead utilizes the cycle lemma from combinatorics.

We mention here one of the applications of (1.1), a 1955 problem of Erdős ([2], [3]) known colloquially as the "multiplication table problem". Let $A(x)$ be the number of positive integers $n \leq x$ which can be written as $n = m_1 m_2$ with each $m_i \leq \sqrt{x}$. Then

$$A(x) \asymp \frac{x}{(\log x)^\delta (\log \log x)^{3/2}}.$$

This follows directly from (1.1) and the inequalities

$$H\left(\frac{x}{4}, \frac{\sqrt{x}}{4}, \frac{\sqrt{x}}{2}\right) \leq A(x) \leq \sum_{k \geq 0} H\left(\frac{x}{2^k}, \frac{\sqrt{x}}{2^{k+1}}, \frac{\sqrt{x}}{2^k}\right).$$

More on the history of estimations of $H(x, y, z)$, further applications and references may be found in [4].

Heuristic argument. For brevity, let $\tau(n, y, z)$ be the number of divisors of n in $(y, z]$. Write $n = n'n''$, where n' is composed only of primes $\leq 2y$ 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 uniformly distributed in $[0, \log n']$. If n' has k prime factors, then the expected value of $\tau(n', y, 2y)$ should be about $\frac{2^k \log 2}{\log n'} \asymp \frac{2^k}{\log y}$. This is $\gg 1$ precisely

---

when $k \geq k_0 + O(1)$, where $k_0 := \left\lfloor \frac{\log \log y}{\log 2} \right\rfloor$. Using the fact (e.g. Theorem 08 of [6]) that the number of $n \leq x$ with n' having k prime factors is of order

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

we obtain a heuristic estimate for $H(x, y, 2y)$ of order

$$\frac{x}{\log y} \sum_{k \geq k_0 + O(1)} \frac{(\log \log y)^k}{k!} \asymp \frac{x(\log \log y)^{k_0}}{k_0! \log y} \asymp \frac{x}{(\log y)^\delta (\log \log y)^{1/2}}.$$

This is slightly too big, and the reason stems from the uniformity assumption about $D(n')$. In fact, for most n' with about k_0 prime factors, the set $D(n')$ is far from uniform, possessing many clusters of divisors and large gaps between clusters. This substantially decreases the likelihood that $\tau(n', y, 2y) \geq 1$. The numbers $\log \log p$ over $p | n'$ are well-known to behave like random numbers in $[0, \log \log 2y]$. Consequently, if we write $n' = p_1 \cdots p_k$, where $p_1 < p_2 < \ldots < p_k$, then we expect $\log \log p_j \approx \frac{j \log \log y}{k_0} = j \log 2 + O(1)$ for each j. Large deviation results from probability theory (see Smirnov's theorem in §4; also see Ch. 1 of [6]) tell us that with high probability there is a j for which $\log \log p_j \leq j \log 2 - c\sqrt{\log \log y}$, where c is a small positive constant. Thus, the 2^j divisors of $p_1 \cdots p_j$ will be clustered in an interval of logarithmic length about $\ll \log p_j \leq 2^j e^{-c\sqrt{\log \log y}}$. On a logarithmic scale, the divisors of n' will then lie in 2^{k-j} translates of this cluster. A measure of the degree of clustering of the divisors of an integer a is given by

$$L(a) = \operatorname{meas} \mathscr{L}(a), \qquad \mathscr{L}(a) = \bigcup_{d | a} [-\log 2 + \log d, \log d).$$

The probability that $\tau(n', y, 2y) \geq 1$ should then be about $L(n')/\log y$. Making this precise leads to the upper and lower bounds for $H(x, y, 2y)$ given below in Lemmas 2.1 and 3.2. The upper bound for $L(a)$ given in Lemma 3.1 (iii) below quantifies how small $L(a)$ must be when there is a j with $\log \log p_j$ considerably smaller than $j \log 2$.

What we really need to count is n for which n' has about k_0 prime factors *and* $L(n') \gg \log n'$. This roughly corresponds to asking for $\log \log p_j \geq j \log 2 - O(1)$ for all j. The anologous problem from statistics theory is to ask for the likelihood than given k_0 random numbers in $[0, 1]$, there are $\leq k_0 x + O(1)$ of them which are $\leq x$, uniformly in $0 \leq x \leq 1$. In section 4, Lemma 4.1, we will see that this probability is about $1/k_0 \asymp 1/\log \log y$ and this leads to the correct order (1.1).

Notation: Let $\tau(n)$ be the number of positive divisors of n, and define $\omega(n)$ to 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 $g \ll f$.

We shall make frequent use of the following estimate, which is a consequence of Mertens' theorem. For a certain constant c_0,

$$(1.2) \qquad \sum_{p \leq x} \frac{1}{p} = \log \log x + c_0 + O(1/\log x) \qquad (x \geq 2).$$

We also need the standard sieve bound (e.g. [5]; Theorem 06 and Exercise 02 of [6])

$$(1.3) \qquad |\{n \le x : P^-(n) > z\}| \asymp \frac{x}{\log z} \qquad (x \ge 2z \ge 4)$$

and Stirling's formula $k! \sim \sqrt{2\pi k}(k/e)^k$.

Remarks/Acknowledgements. This version of the paper incorporates several technical simplifications over the original published version, and (hopefully) an improved exposition. The author has benefited from conversations with a number of people, including Sean Eberhard, Casey Fu, Ben Green, and Dimitrios Koukoulopoulos. In particular, we simplify the upper bound argument using a result on sums of arithmetic functions due to Koukoulopoulos [7, Lemma 2.2], a short proof of which we give below. We also simplify somewhat the lower bound argument using some ideas that grew out of the related paper [1] on permutations with a given sized fixed set.

2. LOWER BOUND

In this section we prove the lower bound implicit in (1.1). The first step is to bound $H(x, y, 2y)$ in terms of a sum of $L(a)/a$. Next, sums of $L(a)/a$ are related via the Cauchy-Schwarz inequality to sums of a function $W(a)$ which counts pairs of divisors of a which are close together. With a strategic choice of sets of a to average over, the problem is reduced to the estimation of a certain combinatorial sum. This is accomplished with the aid of a tool closely related to the so-called "cycle lemma".

Lemma 2.1. *If y_0 is sufficiently large and $y_0 \le y \le \sqrt{x}$, then*

$$H(x, y, 2y) \gg \frac{x}{\log^2 y} \sum_{\substack{P^+(a) \le y^{1/15} \\ \mu^2(a)=1}} \frac{L(a)}{a}.$$

Proof. Consider integers $n = ap_1p_2b \le x$ with p_1 and p_2 prime,

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

and with $\log(y/p_1p_2) \in \mathscr{L}(a)$. The last condition implies that $\tau(ap_1p_2, y, 2y) \ge 1$. In particular, $y^{4/5} \le y/a < p_1p_2 \le 2y$, so that $x/ap_1p_2 \ge x/(2y^{6/5}) \ge \frac{1}{2}y^{4/5}$. Thus, by (1.3), for each triple a, p_1, p_2, the number of possible b is $\gg \frac{x}{ap_1p_2 \log y}$. Now $\mathscr{L}(a)$ is the disjoint union of intervals of length $\ge \log 2$, and for each such interval $[u, v]$ we have

$$\sum_{\substack{u \le \log(y/p_1p_2) \le v \\ y^{1/5} < p_1 < p_2 < \frac{1}{4}y^{4/5}}} \frac{1}{p_1p_2} \ge \sum_{8y^{1/5} < p_1 \le y^{2/5}} \frac{1}{p_1} \sum_{ye^{-v}/p_1 < p_2 \le ye^{-u}/p_1} \frac{1}{p_2} \gg \frac{v - u}{\log y}.$$

Thus, with a fixed, the sum of $\frac{1}{p_1p_2}$ is $\gg \frac{L(a)}{\log y}$ and we obtain

$$H(x, y, 2y) \gg \frac{x}{\log^2 y} \sum_{a \le y^{1/5}} \frac{L(a)}{a}.$$

Next, we replace the sum over a more convenient set, starting with

$$\sum_{a \leq y^{1/5}} \frac{L(a)}{a} \geq \sum_{\substack{a \leq y^{1/5} \\ P^+(a) \leq y^{1/15}, \mu^2(a)=1}} \frac{L(a)}{a} \geq \sum_{\substack{P^+(a) \leq y^{1/15} \\ \mu^2(a)=1}} \frac{L(a)}{a} \left(1 - \frac{\log a}{\log(y^{1/5})}\right).$$

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

$$\sum_{\substack{P^+(a) \leq y^{1/15} \\ \mu^2(a)=1}} \frac{L(a) \log a}{a} = \sum_{\substack{P^+(a) \leq y^{1/15} \\ \mu^2(a)=1}} \frac{L(a)}{a} \sum_{p|a} \log p = \sum_{p \leq y^{1/15}} \frac{\log p}{p} \sum_{\substack{P^+(b) \leq y^{1/15} \\ \mu^2(b)=1, p \nmid b}} \frac{L(pb)}{b}.$$

Using the trivial relation $L(pb) \leq 2L(b)$ and Mertens' estimate, we have

$$\sum_{a \leq y^{1/5}} \frac{L(a)}{a} \geq \sum_{\substack{P^+(a) \leq y^{1/15} \\ \mu^2(a)=1}} \frac{L(a)}{a} \left(1 - \frac{2\log(y^{1/15}) + O(1)}{\log(y^{1/5})}\right) \gg \sum_{\substack{P^+(a) \leq y^{1/15} \\ \mu^2(a)=1}} \frac{L(a)}{a},$$

which concludes the proof of the lemma. □

Lemma 2.2. *For any finite set $\mathscr{A}$ of positive integers,*

$$\sum_{a \in \mathscr{A}} \frac{L(a)}{a} \geq \frac{\left(\sum_{a \in \mathscr{A}} \frac{\tau(a)}{a}\right)^2}{6 \sum_{a \in \mathscr{A}} \frac{W(a)}{a}},$$

where

$$W(a) = |\{(d, d') : d|a, d'|a, |\log d/d'| \leq \log 2\}|.$$

Proof. Since $\tau(a) \log 2 = \int \tau(a, e^u, 2e^u) \, du$, by the Cauchy-Schwarz inequality,

$$\left(\sum_{a \in \mathscr{A}} \frac{\tau(a)}{a}\right)^2 (\log 2)^2 = \left(\sum_{a \in \mathscr{A}} \frac{1}{a} \int \tau(a, e^u, 2e^u) \, du\right)^2$$

$$\leq \left(\sum_{a \in \mathscr{A}} \frac{L(a)}{a}\right) \left(\sum_{a \in \mathscr{A}} \frac{1}{a} \int \tau^2(a, e^u, 2e^u) \, du\right).$$

Let $k_j = \tau(a, 2^{j-1}, 2^j)$ for each integer j. Then

$$\int \tau^2(a, e^u, 2e^u) \, du \leq (\log 2) \sum_j (k_j + k_{j+1})^2 \leq 4(\log 2) \sum_j k_j^2 \leq 4(\log 2) W(a). \quad \square$$

We apply Lemma 2.2 with sets $\mathscr{A}$ of integers whose prime factors are localized. To simplify later analysis, partition the primes into sets $D_1, D_2, \ldots$, 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

(2.1) $$\sum_{\lambda_{j-1} < p \leq \lambda_j} \frac{1}{p} \leq \log 2.$$

For example, $\lambda_1 = 2$ and $\lambda_2 = 7$. By (1.2), we have

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

and it follows that for some absolute constant K,

$$(2.2) \qquad 2^{j-K} \le \log \lambda_j \le 2^{j+K} \qquad (j \ge 0).$$

For a vector $\mathbf{b} = (b_1, \ldots, b_J)$ of non-negative integers, let $\mathscr{A}(\mathbf{b})$ be the set of square-free integers a composed of exactly b_j prime factors from D_j for each j.

Lemma 2.3. *Assume* $\mathbf{b} = (b_1, \ldots, b_J)$. *Then*

$$\sum_{a \in \mathscr{A}(\mathbf{b})} \frac{W(a)}{a} \ll \frac{(2 \log 2)^{b_1 + \cdots + b_J}}{b_1! \cdots b_J!} \sum_{j=1}^{J} 2^{-j + b_1 + \cdots + b_j}.$$

Proof. Let $B = b_1 + \cdots + b_J$ and for $j \ge 0$ let $B_j = \sum_{i \le j} b_i$. Let $a = p_1 \cdots p_B$, where

$$(2.3) \qquad p_{B_{j-1}+1}, \ldots, p_{B_j} \in D_j \qquad (1 \le j \le J)$$

and the primes in each interval D_j are unordered. Since $W(p_1 \cdots p_B)$ is the number of pairs $Y, Z \subseteq \{1, \ldots, B\}$ with

$$(2.4) \qquad \left| \sum_{i \in Y} \log p_i - \sum_{i \in Z} \log p_i \right| \le \log 2,$$

we have

$$(2.5) \qquad \sum_{a \in \mathscr{A}(\mathbf{b})} \frac{W(a)}{a} \le \frac{1}{b_1! \cdots b_J!} \sum_{Y, Z \subseteq \{1, \ldots, B\}} \sum_{\substack{p_1, \ldots, p_B \\ (2.3), (2.4)}} \frac{1}{p_1 \cdots p_B}.$$

When $Y = Z$, (2.1) implies that the inner sum on the right side of (2.5) is $\le (\log 2)^B$, and there are 2^B such pairs Y, Z. When $Y \ne Z$, let $I = \max[(Y \cup Z) - (Y \cap Z)]$. With all the p_i fixed except for p_I, (2.4) implies that $U \le p_I \le 4U$ for some number U. Let $E(I)$ be defined by $B_{E(I)-1} < I \le B_{E(I)}$, i.e. $p_I \in D_{E(I)}$. By (1.2),

$$\sum_{\substack{U \le p_I \le 4U \\ p_I \in D_{E(I)}}} \frac{1}{p_I} \ll \frac{1}{\max(\log U, \log \lambda_{E(I)-1})} \ll 2^{-E(I)}.$$

Thus, by (2.1) the inner sum in (2.5) is $\ll 2^{-E(I)}(\log 2)^B$. With I fixed, there correspond $2^{B-I+1} 4^{I-1} = 2^{B+I-1}$ pairs Y, Z. By (2.5),

$$\sum_{a \in \mathscr{A}(\mathbf{b})} \frac{W(a)}{a} \ll \frac{(2 \log 2)^B}{b_1! \cdots b_J!} \left[1 + \sum_{I=1}^{B} 2^{I-E(I)} \right] \ll \frac{(2 \log 2)^B}{b_1! \cdots b_J!} \sum_{j=1}^{J} 2^{-j} \sum_{B_{j-1} < I \le B_j} 2^I,$$

and the claimed bound follows. $\qquad \square$

Now suppose that M is a sufficiently large, fixed positive integer, $b_i = 0$ for $i < M$, and $b_j \le Mj$ for each j. By (2.2),

$$\sum_{a \in \mathscr{A}(\mathbf{b})} \frac{\tau(a)}{a} = 2^k \prod_{j=M}^{J} \frac{1}{b_j!} \left(\sum_{p_1 \in D_j} \frac{1}{p_1} \sum_{\substack{p_2 \in D_j \\ p_2 \ne p_1}} \frac{1}{p_2} \cdots \sum_{\substack{p_{b_j} \in D_j \\ p_{b_j} \notin \{p_1, \ldots, p_{b_j-1}\}}} \frac{1}{p_{b_j}} \right)$$

$$(2.6)$$

$$\ge 2^k \prod_{j=M}^{J} \frac{1}{b_j!} \left(\log 2 - \frac{b_j}{\lambda_{j-1}} \right)^{b_j} \ge \frac{(2 \log 2)^k}{2 b_M! \cdots b_J!}.$$

Let

$$k = \left\lfloor \frac{\log \log y}{\log 2} - 2M \right\rfloor, \qquad J = M + k - 1.$$

Let $\mathscr{B}$ be the set of vectors $(b_1, \ldots, b_J)$ with $b_i = 0$ for $i < M$ and $b_1 + \cdots + b_J = k$. Let $\mathscr{B}^*$ be the set of $\mathbf{b} \in \mathscr{B}$ with $b_j \leq Mj$ for each $j \geq M$. If $\mathbf{b} \in \mathscr{B}$ and $a \in \mathscr{A}(\mathbf{b})$, then by (2.2), $P^+(a) \leq y^{1/15}$ since $2^{J+1} \leq 2^{-M} \log y$. Put

$$(2.7) \qquad f(\mathbf{b}) = \sum_{h=M}^{J} 2^{M-1-h+b_M+\cdots+b_h}.$$

By Lemma 2.3,

$$\sum_{a \in \mathscr{A}(\mathbf{b})} \frac{W(a)}{a} \ll \frac{(2 \log 2)^k}{b_M! \cdots b_J!} \left(1 + 2^{1-M} f(\mathbf{b})\right) \ll \frac{(2 \log 2)^k}{b_M! \cdots b_J!} f(\mathbf{b})$$

since $f(\mathbf{b}) \geq 1/2$. By Lemmas 2.1 and 2.2, plus (2.6), we have for large y

$$(2.8) \qquad H(x, y, 2y) \gg \frac{x(2 \log 2)^k}{\log^2 y} \sum_{\mathbf{b} \in \mathscr{B}^*} \frac{1}{b_M! \cdots b_J! f(\mathbf{b})}.$$

Observe that the product of factorials is unchanged under permutation of $b_M, \ldots, b_J$. Roughly speaking,

$$f(\mathbf{b}) \approx g(\mathbf{b}) := \max_j 2^{(b_M-1)+\cdots+(b_j-1)}.$$

Note that $(b_M - 1) + \cdots + (b_J - 1) = k - (J - M + 1) = 0$.

Given real numbers $z_1, \cdots, z_k$ with zero sum, there is a cyclic permutation $\mathbf{z}'$ of the vector $\mathbf{z} = (z_1, \ldots, z_k)$ all of whose partial sums are ≥ 0: let i be the index minimizing $z_1 + \cdots + z_i$ and take $\mathbf{z}' = (z_{i+1}, \ldots, z_k, z_1, \ldots, z_i)$. In combinatorics, this fact is know as the *cycle lemma*. Thus, there is a a cyclic permutation $\mathbf{b}'$ of $\mathbf{b}$ with $g(\mathbf{b}') = 1$. Thus, we expect that $1/f(\mathbf{b}')$ will be $\gg 1/k$ on average over $\mathbf{b}'$ and that $1/f(\mathbf{b}) \gg 1/k$ on average over $\mathbf{b} \in \mathscr{B}$. This is essentially what we prove next; see (2.10) below.

Lemma 2.4. *For positive real numbers $x_1, \ldots, x_r$ with product X, let $x_{r+i} = x_i$ for $i \geq 1$. Then*

$$\sum_{j=0}^{r-1} \left(\sum_{h=1}^{r} x_{1+j} \cdots x_{h+j} \right)^{-1} \in \left[\frac{1}{\max(1, X)}, \frac{1}{\min(1, X)} \right].$$

Proof. Put $y_0 = 1$ and $y_j = x_1 \cdots x_j$ for $j \geq 1$. The sum in question is

$$\sum_{j=0}^{r-1} \left(\sum_{h=1}^{r} \frac{y_{h+j}}{y_j} \right)^{-1} = \sum_{j=0}^{r-1} \frac{y_j}{y_{1+j} + \cdots + y_{r+j}}.$$

Since $y_r = X$,

$$y_{1+j} + \cdots + y_{r+j} = X(y_0 + \cdots + y_j) + y_{1+j} + \cdots + y_{r-1}$$
$$\in [\min(1, X)(y_0 + \cdots + y_{r-1}), \max(1, X)(y_0 + \cdots + y_{r-1})]. \qquad \square$$

We have

$$(2.9) \qquad \sum_{\mathbf{b} \in \mathscr{B}^*} \frac{1}{b_M! \cdots b_J! f(\mathbf{b})} \geq S_0 - \sum_{M \leq j < k/M} S_1(j),$$

where

$$S_0 = \sum_{\mathbf{b}\in\mathscr{B}} \frac{1}{b_M!\cdots b_J!f(\mathbf{b})}, \qquad S_1(j) = \sum_{\substack{\mathbf{b}\in\mathscr{B}\\ b_j>Mj}} \frac{1}{b_M!\cdots b_J!f(\mathbf{b})}.$$

Let $x_i = 2^{-1+b_{M-1+i}}$ for $1 \le i \le k$. Then $x_1\cdots x_k = 1$ and

$$f(\mathbf{b}) = x_1 + x_1 x_2 + \cdots + x_1 x_2 \cdots x_k.$$

By Lemma 2.4 and the multinomial theorem,

$$(2.10) \qquad S_0 = \sum_{\mathbf{b}\in\mathscr{B}} \frac{1}{b_M!\cdots b_J!} \frac{1}{k} \sum_{j=0}^{k-1} \left(\sum_{h=1}^{k} x_{1+j}\cdots x_{h+j} \right)^{-1} = \frac{k^{k-1}}{k!}.$$

To bound $S_1(j)$, apply Lemma 2.4 with $x_i = 2^{b_{j+i}-1}$ for $1 \le i \le J - j$ and note that

$$X = x_1 \cdots x_{J-j} = 2^{j+1-M-b_M-\cdots-b_j} < 1.$$

From the definition of J and our assumption $j \le k/M$, we have $J - j \ge k/2$. Write $\mathbf{b}' = (b_M, \ldots, b_{j-1}, b_{j+1}, \ldots, b_J)$, whose sum of components is $k - b_j$. Ignoring the terms with $h \le j$ in (2.7), using Lemma 2.4 and the multinomial theorem, we find

$$\begin{aligned}
S_1(j) &\le \sum_{\substack{b_j>Mj}} \frac{1}{b_j!} \sum_{\mathbf{b}'} \frac{1}{\prod_{i\ne j} b_i!} \frac{1}{2^{M-1-j+b_M+\cdots+b_j}} \frac{1}{J-j} \sum_{i=0}^{J-j-1} \left(\sum_{h=1}^{J-j} x_{1+i}\cdots x_{h+i} \right)^{-1} \\
&\le \frac{1}{J-j} \sum_{b_j>Mj} \frac{(k-1)^{k-b_j}}{b_j!(k-b_j)!} \\
&\le \frac{2}{k} \sum_{b_j>Mj} \frac{(k-1)^k k(k-1)\cdots(k-b_j+1)}{b_j!k!(k-1)^{b_j}} \\
&= \frac{2k^{k-1}}{k!} \sum_{b_j>Mj} \frac{(k-1)\cdots(k-b_j+1)}{(k-1)^{b_j-1}b_j!} \\
&\le \frac{2k^{k-1}}{k!} \sum_{b_j>Mj} \frac{1}{b_j!} \le \frac{k^{k-1}}{k!} \frac{2}{(Mj)!}.
\end{aligned}$$

Hence, if $M \ge 2$ then

$$(2.11) \qquad \sum_{M\le j<k/M} S_1(j) \le \frac{k^{k-1}}{10k!}.$$

By (2.9), (2.10), and (2.11),

$$\sum_{\mathbf{b}\in\mathscr{B}^*} \frac{1}{b_M!\cdots b_J!f(\mathbf{b})} \ge \frac{k^{k-1}}{2k!}.$$

The lower bound in (1.1) for large y now follows from (2.8) and Stirling's formula. If $y \le y_0$ for some fixed constant y_0, the lower bound in (1.1) follows from $H(x, y, 2y) \gg x$.

3. UPPER BOUND, PART I

In this section, we prove the upper bound implicit in (1.1), except for the estimation of some integrals which will be dealt with in section 4. As with the lower bound argument, we begin by bounding $H(x, y, 2y)$ in terms of a sum involving $L(a)$. Using a relatively simple upper bound for $L(a)$ proved in Lemma 3.1 below, the sums involving $L(a)$ are bounded in terms of particular multivariate integrals. The estimates for these integrals in section 4 allow us then to complete the proof.

Lemma 3.1. *We have*

   (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* $p_1 < \cdots < p_k$, *then*

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

Proof. Part (i) is immediate, since $\mathscr{L}(a)$ is the union of $\tau(a)$ intervals of length $\log 2$, all contained in $[-\log 2, \log a)$. Part (ii) follows from

$$\mathscr{L}(ab) = \bigcup_{d|b} \{u + \log d : u \in \mathscr{L}(a)\}.$$

Combining parts (i) and (ii) with $a = p_1 \cdots p_j$ and $b = p_{j+1} \cdots p_k$ yields (iii). □

Introduce the notation

$$\mathscr{P}(x) = \{n \in \mathbb{N} : \mu^2(n) = 1, P^+(n) \leq x\}.$$

Lemma 3.2. *If* $3 \leq y \leq \sqrt{x}$, *then*

$$H(x, y, 2y) \ll x \sum_{a \in \mathscr{P}(2y)} \frac{L(a)}{a \log^2(y^{2/3}/a + P^+(a))}.$$

Proof. We may assume that y is sufficiently large (say $y \geq y_0$), as the statement is trivial for small y. First, we relate $H(x, y, 2y)$ to $H^*(x, y, z)$, the number of *squarefree* integers $n \leq x$ with $\tau(n, y, z) \geq 1$. Write $n = n'n''$, where n' is squarefree, n'' is squarefull and $(n', n'') = 1$. The number of $n \leq x$ with $n'' > y^{1/10}$ is

$$\leq x \sum_{n'' > y^{1/10}} \frac{1}{n''} \ll \frac{x}{y^{1/20}}.$$

If $n'' \leq y^{1/10}$, then for some $f|n''$, n' has a divisor in $(y/f, 2y/f]$, hence

$$(3.1) \qquad H(x, y, 2y) \leq \sum_{n'' \leq y^{1/10}} \sum_{f|n''} H^* \left(\frac{x}{n''}, \frac{y}{f}, \frac{2y}{f} \right) + O\left(\frac{x}{y^{1/20}} \right).$$

Next, we show that for $y_0 \leq y_1 \leq x_1^{5/9}$,

$$(3.2) \qquad H^*(x_1, y_1, 2y_1) - H^*(\tfrac{1}{2} x_1, y_1, 2y_1) \ll x_1 \sum_{a \in \mathscr{P}(2y_1)} \frac{L(a)}{a \log^2 \left(y_1^{4/5}/a + P^+(a) \right)}.$$

Let $\mathscr{A}$ be the set of squarefree integers $n \in (\frac{1}{2} x_1, x_1]$ with a divisor in $(y_1, 2y_1]$. Put $z_1 = 2y_1$, $y_2 = \frac{x_1}{4y_1}$, $z_2 = \frac{x_1}{y_1}$. If $n \in \mathscr{A}$, 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 $p = P^+(m_j) < P^+(m_{3-j})$. Write $n = abp$, where $P^+(a) < p < P^-(b)$ and $b > p$. Since $\tau(ap, y_j, z_j) \geq 1$, we have $p \geq y_j/a \geq y_1^{4/5}/a$. We also have $p \leq \min(z_1, z_2) \leq 2y_1$. By (1.3), given a and p, the number of possible b is

$$\ll \frac{x_1}{ap \log p} \leq \frac{x_1}{ap \log \max \left(P^+(a), y_1^{4/5}/a \right)},$$

Since a has a divisor in $(y_j/p, z_j/p]$, we have $\log(y_j/p) \in \mathscr{L}(a)$ or $\log(2y_j/p) \in \mathscr{L}(a)$. Since $\mathscr{L}(a)$ is the disjoint union of intervals of length $\geq \log 2$ with total measure $L(a)$, by repeated use of (1.2) we obtain

$$\sum_{\substack{\log(cy_j/p) \in \mathscr{L}(a) \\ p \geq P^+(a)}} \frac{1}{p} \ll \frac{L(a)}{\log \max \left(P^+(a), y_1^{4/5}/a \right)} \qquad (c = 1, 2),$$

and (3.2) follows.

Write $x_2 = x/n''$, $y_1 = y/f$. Each $n \in (x_2/y_1^{1/10}, x_2]$ lies in an interval $(2^{-r+1}x_2, 2^{-r}x_2]$ for some integer $0 \leq r \leq \frac{\log y_1}{10 \log 2}$. Applying (3.2) with $x_1 = 2^{-r}x_2$ for each r gives

$$H^*(x_2, y_1, 2y_1) \ll \frac{x_2}{y_1^{1/10}} + x_2 \sum_{a \in \mathscr{P}(2y_1)} \frac{L(a)}{a \log^2 \left(y_1^{4/5}/a + P^+(a) \right)}.$$

The first term $x_2/y_1^{1/10}$ may be ignored because $L(1) \gg 1$. Thus, by (3.1),

$$H(x, y, 2y) \ll \frac{x}{y^{1/20}} + x \sum_{n'' \leq y^{1/10}} \frac{1}{n''} \sum_{f | n''} \sum_{a \in \mathscr{P}(y/f)} \frac{L(a)}{a \log^2 ((y/f)^{4/5}/a + P^+(a))}.$$

Again, the term $x/y^{1/20}$ is negligible. We have $(y/f)^{4/5} \geq y^{2/3}$ for any pair (n'', f) and $\sum_{n''} \tau(n'')/n'' \ll 1$. This completes the proof. □

The next lemma is a special case of a result of Koukoulopoulos [7, Lemma 2.2].

Lemma 3.3. *We have*

$$\sum_{a \in \mathscr{P}(2y)} \frac{L(a)}{a \log^2(P^+(a) + y^{2/3}/a)} \ll \frac{1}{(\log y)^2} \sum_{a \in \mathscr{P}(2y)} \frac{L(a)}{a}.$$

Proof. Let $\mathscr{P}_1 = \{a \in \mathscr{P}(y^{1/4}) : a > y^{1/2}\}$. Then clearly

$$\sum_{a \in \mathscr{P}(2y)} \frac{L(a)}{a \log^2(P^+(a) + y^{2/3}/a)} \ll \sum_{a \in \mathscr{P}_1} \frac{L(a)}{a \log^2 P^+(a)} + \frac{1}{(\log y)^2} \sum_{a \in \mathscr{P}(2y)} \frac{L(a)}{a}.$$

For $a \in \mathscr{P}_1$, let $p = P^+(a)$ and $a = pb$, so $b > y^{1/4}$. By Lemma 3.1 (ii), $L(ab) \leq 2L(b)$ and thus

$$\sum_{a \in \mathscr{P}_1} \frac{L(a)}{a \log^2 P^+(a)} \leq 2 \sum_{p \leq y^{1/4}} \frac{1}{p \log^2 p} \sum_{\substack{b \in \mathscr{P}(p) \\ b > y^{1/4}}} \frac{L(b)}{b}$$

$$\leq 2 \sum_{p \leq y^{1/4}} \frac{1}{p \log^2 p} \frac{4^3}{\log^3 y} \sum_{b \in \mathscr{P}(p)} \frac{L(b) \log^3 b}{b}.$$

Next,

$$\sum_{b \in \mathscr{P}(p)} \frac{L(b) \log^3 b}{b} = \sum_{b \in \mathscr{P}(p)} \frac{L(b)}{b} \sum_{p_1|b, p_2|b, p_3|b} (\log p_1)(\log p_2)(\log p_3)$$

$$\leq 8 \sum_{p_1, p_2, p_3 \leq p} \frac{(\log p_1)(\log p_2)(\log p_3)}{[p_1, p_2, p_3]} \sum_{t \in \mathscr{P}(p)} \frac{L(t)}{t},$$

where we have written $b = [p_1, p_2, p_3]t$. Considering separately the three cases ($p_1 = p_2 = p_3$, two of the p_i equal, all p_i distinct), we find that

$$\sum_{p_1, p_2, p_3 \leq p} \frac{(\log p_1)(\log p_2)(\log p_3)}{[p_1, p_2, p_3]} \ll (\log p)^3$$

by Mertens' estimate. Extending the range of t to $t \in \mathscr{P}(2y)$, we get

$$\sum_{a \in \mathscr{P}_1} \frac{L(a)}{a \log^2(P^+(a) + y^{2/3}/a)} \ll \sum_{p \leq y^{1/4}} \frac{1}{p \log^2 p} \frac{(\log p)^3}{(\log y)^3} \sum_{t \in \mathscr{P}(2y)} \frac{L(t)}{t}.$$

A final application of Mertens' estimate concludes the proof. $\square$

Combining Lemma 3.2 with 3.3 produces an upper bound of the same shape as the lower bound in Lemma 2.1.

Lemma 3.4. *Uniformly for $3 \leq y \leq \sqrt{x}$, we have*

$$H(x, y, 2y) \ll \frac{x}{\log^2 y} \sum_{a \in \mathscr{P}(2y)} \frac{L(a)}{a}.$$

We cut up the sum in Lemma 3.4 according to $\omega(a)$. Let

$$T_k(y) = \sum_{\substack{a \in \mathscr{P}(2y) \\ \omega(a) = k}} \frac{L(a)}{a}.$$

We next bound $T_k(y)$ in terms of a mutivariate integral. Since $\sum_{p \leq z} 1/p = \log \log z + O(1)$, by partial summation we expect for "nice" functions f that

$$\sum_{p_1 < \cdots < p_k \leq P} \frac{f\left(\frac{\log \log p_1}{\log \log P}, \cdots, \frac{\log \log p_k}{\log \log P}\right)}{p_1 \cdots p_k} \approx (\log \log P)^k \int \cdots \int_{0 \leq \xi_1 \leq \cdots \leq \xi_k \leq 1} f(\boldsymbol{\xi}) \, d\boldsymbol{\xi}.$$

For $a = p_1 \ldots p_k$, the function $L(a)$ is not very regular as a function of $p_1, \ldots, p_k$. However, the most common way for $L(a)$ to be small is for a to have many small prime factors, and the bound in Lemma 3.1 (iii) captures this nicely. Moreover, this bound has the useful property of being monotone in each variable.

Lemma 3.5. *Suppose y is large, $v = \left\lfloor \frac{\log \log y}{\log 2} \right\rfloor$ and $1 \leq k \leq 10v$. Then*

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

Proof. Recall the definition of λ_i, D_i from §2. Consider $a = p_1 \cdots p_k$, where $p_1, \ldots, p_k \leq 2y$, the primes p_i are unordered and define j_i by $p_i \in D_{j_i}$ $(1 \leq i \leq k)$. By (2.2), $0 \leq j_i \leq v + K + 2$ for each i. Let $\tilde{p}_1 \leq \tilde{p}_2 \leq \cdots$ be the increasing rearrangement of $p_1, p_2, \ldots$ and let $\tilde{j}_1 \leq \tilde{j}_2 \leq \cdots$ be the increasing rearrangement of $j_1, j_2, \ldots$. In particular, $\tilde{p}_i \in D_{\tilde{j}_i}$ for each i. By Lemma 3.1 (iii) and (2.2),

$$L(a) \leq 2^k \min_{0 \leq g \leq k} 2^{-g}(\log(2\tilde{p}_1 \cdots \tilde{p}_g)) \leq 2^{k+K} F(\mathbf{j}),$$

where

$$F(\mathbf{j}) = \min_{0 \leq g \leq k} 2^{-g}(2^{\tilde{j}_1} + \cdots + 2^{\tilde{j}_g} + 1).$$

Then

$$T_k(y) \leq \frac{2^{k+K}}{k!} \sum_{j_1, \ldots, j_k = 1}^{v+K+2} F(\mathbf{j}) \sum_{\substack{p_1, \ldots, p_k \\ p_i \in D_{j_i} \ (1 \leq i \leq k)}} \frac{1}{p_1 \cdots p_k} \leq \frac{2^K}{k!} (2 \log 2)^k \sum_{\mathbf{j}} F(\mathbf{j}).$$

Extend the domain of F to include k-tuples of positive real numbers. It is clear that if $j_i \leq t_i \leq j_i + 1$ for each i, and $\tilde{t}_1 \leq \tilde{t}_2 \leq \cdots$ is the increasing rearrangement of $t_1, t_2, \ldots$, then $\tilde{t}_i \geq \tilde{j}_i$ for every i. Consequently, $F(\mathbf{j}) \leq F(\mathbf{t})$. Therefore,

$$\sum_{\mathbf{j}} F(\mathbf{j}) = \sum_{\mathbf{j}} F(\mathbf{j}) \int_{j_1}^{j_1+1} \cdots \int_{j_k}^{j_k+1} 1 d\mathbf{t} \leq \int_0^{v+K+3} \cdots \int_0^{v+K+3} F(\mathbf{t}) \, d\mathbf{t}$$

Observe that $F(\mathbf{t})$ is symmetric in $t_1, \ldots, t_k$. Making the change of variables $t_i = (v+K+3)\xi_i$ for each i, we see that the multiple integral on the right side equals

$$(v + K + 3)^k k! \int \cdots \int_{0 \leq \xi_1 \leq \cdots \leq \xi_k \leq 1} \min_{0 \leq g \leq k} 2^{-g} \left(2^{(v+K+3)\xi_i} + \cdots + 2^{(v+K+3)\xi_g} + 1 \right) d\boldsymbol{\xi}.$$

We conclude that

$$T_k(y) \leq 2^{2K+3}((2\log 2)(v + K + 3))^k U_k(v).$$

Lastly, $((2\log 2)(v + K + 3))^k \ll (2\log\log y)^k$ since $k \leq 10v$, and the lemma follows. $\qquad\square$

Estimating $U_k(v)$ is the most complex part of the argument. The next lemma will be proved in section 4.

Lemma 3.6. *Suppose k, v are integers with $0 \leq k \leq 10v$. Then*

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

Notice that the bound in Lemma 3.6 undergoes a change of behavior at $k = v$.

Proof of (1.1), *upper bound.* Let $v = \left\lfloor \frac{\log\log y}{\log 2} \right\rfloor$. By Lemmas 3.5 and 3.6,

$$\sum_{v \leq k \leq 10v} T_k(y) \ll \sum_{v \leq k \leq 10v} \frac{(k-v)^2 + 1}{2^{k-v}} \frac{(2\log\log y)^k}{(k+1)!} \ll \frac{(2\log\log y)^v}{(v+1)!}$$

and

$$\sum_{1 \leq k \leq v} T_k(y) \ll \sum_{1 \leq k \leq v} \frac{((v-k)^2 + 1)(2\log\log y)^k}{(k+1)!} \ll \frac{(2\log\log y)^v}{(v+1)!}.$$

By Lemma 3.1 (i),

$$\sum_{k \geq 10v} T_k(y) \leq \sum_{k \geq 10v} \sum_{\substack{a \in \mathscr{P}(2y) \\ \omega(a)=k}} \frac{2^k \log 2}{a} \leq \sum_{k \geq 10v} \frac{2^k}{k!} \left(\sum_{p \leq 2y} \frac{1}{p} \right)^k$$

$$\leq \frac{(2 \log \log y + O(1))^{10v}}{(10v)!} \ll \frac{(2 \log \log y)^v}{(v+1)!}.$$

Finally, $T_0(y) = L(1) = \log 2$. Recalling the definition of v and combining the above bounds on $T_k(y)$ with Stirling's formula and Lemma 3.4 completes the proof. $\square$

4. UPPER BOUND, PART II

The goal of this section is to prove Lemma 3.6, and thus complete the proof of the upper bound in (1.1).

Let $Y_1, \ldots, Y_n$ be independent, uniformly distributed random variables in $[0,1]$. Let ξ_1 be the smallest of the numbers Y_i, let ξ_2 be the next smallest, etc., so that $0 \leq \xi_1 \leq \cdots \leq \xi_n \leq 1$. The numbers ξ_i are the *order statistics* for $Y_1, \ldots, Y_n$. Then $k!U_k(v)$ is the expectation of the random variable

$$X = \min_{0 \leq j \leq k} 2^{-j}(2^{v\xi_1} + \cdots + 2^{v\xi_j} + 1).$$

Heuristically, we expect that

$$(4.1) \qquad \mathbb{E}X \ll \mathbb{E} \min_{1 \leq j \leq k} 2^{-j+v\xi_j},$$

so we need to understand the distribution of $\min_{1 \leq j \leq k} v\xi_j - j$. Let $Q_k(u,v)$ be the probability that $\xi_i \geq \frac{i-u}{v}$ for every i. In the special case $v = k$, Smirnov in 1939 showed that

$$Q_k(x\sqrt{k}, k) \sim 1 - e^{-2x^2}$$

for each fixed x. The corresponding probability estimate for two-sided bounds on the ξ_i was established by Kolmogorov in 1933 and together these limit theorems are the basis of the *Kolmogorov-Smirnov goodness-of-fit statistical tests*.

In the next lemma, we prove new, uniform estimates for $Q_k(u,v)$. The remainder of the section is essentially devoted to proving (4.1). The details are complicated, but the basic idea is that if $2^{-j}(2^{v\xi_1} + \cdots + 2^{v\xi_j})$ is much large than $2^{v\xi_j-j}$, then for some large l, the numbers $\xi_{j-l}, \ldots, \xi_j$ are all very close to one another. As shown below in Lemmas 4.3 and 4.4, this is quite rare.

Lemma 4.1. *Let $w = u + v - k$. Uniformly in $u \geq 0$ and $w \geq 0$, we have*

$$Q_k(u,v) \ll \frac{(u+1)(w+1)^2}{k}.$$

Proof. Without loss of generality, suppose $k \geq 100$, $u \leq k/10$ and $w \leq \sqrt{k}$. If $\min_{1 \leq i \leq k}(\xi_i - \frac{i-u}{v}) < 0$, let l be the smallest index with $\xi_l < \frac{l-u}{v}$ and write $\xi_l = \frac{l-u-\lambda}{v}$, so that $0 \leq \lambda \leq 1$. Let

$$R_l(\lambda) = \text{Vol} \left\{ 0 \leq \xi_1 \leq \cdots \leq \xi_{l-1} \leq \frac{l-u-\lambda}{v} : \xi_i \geq \frac{i-u}{v} \ (1 \leq i \leq l-1) \right\}.$$

Then we have

$$Q_k(u,v) = 1 - \frac{k!}{v} \int_0^1 \sum_{u+\lambda \leq l \leq k} R_l(\lambda) \operatorname{Vol}\left\{\frac{l-u-\lambda}{v} \leq \xi_{l+1} \leq \cdots \leq \xi_k \leq 1\right\} d\lambda$$

$$= 1 - \frac{k!}{v} \int_0^1 \sum_{u+\lambda \leq l \leq k} \frac{R_l(\lambda)}{(k-l)!} \left(\frac{k+w+\lambda-l}{v}\right)^{k-l} d\lambda.$$

Now suppose that $\xi_k \leq 1 - \frac{2w+2}{v} = \frac{k-u-w-2}{v}$. Then $\min_{1 \leq i \leq k}(\xi_i - \frac{i-u}{v}) < 0$. Defining l and λ as before, we have

$$\left(1 - \frac{2w+2}{v}\right)^k = k! \operatorname{Vol}\left\{0 \leq \xi_1 \leq \cdots \leq \xi_k \leq 1 - \frac{2w+2}{v}\right\}$$

$$= \frac{k!}{v} \int_0^1 \sum_{u+\lambda \leq l \leq k-w-2+\lambda} \frac{R_l(\lambda)}{(k-l)!} \left(\frac{k-l-w-2+\lambda}{v}\right)^{k-l} d\lambda.$$

Thus, for any $A > 0$, we have

$$Q_k(u,v) = 1 - A\left(1 - \frac{2w+2}{v}\right)^k - \frac{k!}{v} \int_0^1 \sum_{k-w-2+\lambda < l \leq k} \frac{R_l(\lambda)}{(k-l)!} \left(\frac{k+w+\lambda-l}{v}\right)^{k-l} d\lambda$$

$$+ \frac{k!}{v} \int_0^1 \sum_{u+\lambda \leq l \leq k-w-2+\lambda} \frac{R_l(\lambda)}{(k-l)! v^{k-l}} \left[A(k-l-w-2+\lambda)^{k-l} - (k-l+w+\lambda)^{k-l}\right] d\lambda.$$

Noting that $2 - \lambda \geq \lambda$, we have

$$\left(\frac{k-l-w-2+\lambda}{k-l+w+\lambda}\right)^{k-l} = \left(1 - \frac{w+2-\lambda}{k-l}\right)^{k-l} \left(1 + \frac{w+\lambda}{k-l}\right)^{-(k-l)}$$

$$= \exp\left\{-(2w+2) + \sum_{j=2}^{\infty} \frac{-(w+2-\lambda)^j + (-1)^j(w+\lambda)^j}{j(k-l)^{j-1}}\right\}$$

$$\leq e^{-(2w+2)}.$$

Thus, taking $A = e^{2w+2}$, we conclude that

$$Q_k(u,v) \leq 1 - e^{2w+2}\left(1 - \frac{2w+2}{v}\right)^k$$

$$= 1 - \exp\left\{\frac{2w+2}{v}(v - k + O(w))\right\}$$

$$= 1 - \exp\left\{\frac{-2uw + O(u+w^2+1)}{v}\right\}$$

$$\leq \frac{2uw + O(u+w^2+1)}{v} \ll \frac{(u+1)(w+1)^2}{k}. \qquad \square$$

Lemma 4.2. *If $t \geq 2$, $b \geq 0$ and $t+a+b > 0$, then*

$$\sum_{\substack{1 \leq j \leq t-1 \\ j+a>0}} \binom{t}{j} (a+j)^{j-1}(b+t-j)^{t-j-1} \leq e^4(t+a+b)^{t-1}.$$

Proof. Let $C_t(a, b)$ denote the sum in the lemma. We may assume that $a > 1 - t$, otherwise $C_t(a, b) = 0$. The associated "complete" sum is evaluated exactly using one of Abel's identities ([8], p.20, equation (20))

$$(4.2) \qquad \sum_{j=0}^{t} \binom{t}{j}(a + j)^{j-1}(b + t - j)^{t-j-1} = \left(\frac{1}{a} + \frac{1}{b}\right)(t + a + b)^{t-1} \qquad (ab \neq 0).$$

If $a \geq -1$, put $A = \max(1, a)$ and $B = \max(1, b)$. By (4.2),

$$
\begin{aligned}
C_t(a, b) \leq C_t(A, B) &\leq \left(\frac{1}{A} + \frac{1}{B}\right)(t + A + B)^{t-1} \\
(4.3) \qquad\qquad &\leq 2(t + a + b + 3)^{t-1} \\
&\leq 2e^{\frac{3(t-1)}{t+a+b}}(t + a + b)^{t-1} < e^4(t + a + b)^{t-1}.
\end{aligned}
$$

Next assume $a < -1$. Since $(1 + c/x)^x$ is an increasing function for $x > 1$, we have

$$(a + j)^{j-1} = (j - 1)^{j-1}\left(1 + \frac{a+1}{j-1}\right)^{j-1} \leq (j - 1)^{j-1}\left(1 + \frac{a+1}{t-1}\right)^{t-1}.$$

Thus, by (4.3),

$$
\begin{aligned}
C_t(a, b) &\leq \left(\frac{t + a}{t - 1}\right)^{t-1} C_t(-1, b) \\
&\leq e^4 \left(\frac{(t + a)(t + b - 1)}{t - 1}\right)^{t-1} = e^4\left(t + a + b + \frac{(a+1)b}{t-1}\right)^{t-1} \\
&\leq e^4(t + a + b)^{t-1}. \qquad\qquad\qquad\qquad\qquad\qquad\qquad \square
\end{aligned}
$$

For brevity, write

$$S_k(u, v) = \left\{\boldsymbol{\xi} : 0 \leq \xi_1 \leq \cdots \leq \xi_k \leq 1 : \xi_i \geq \frac{i - u}{v}\ (1 \leq i \leq k)\right\},$$

so that $Q_k(u, v) = k! \operatorname{Vol} S_k(u, v)$.

Lemma 4.3. *Suppose $g, k, s, u, v \in \mathbb{Z}$ satisfy*

$$1 \leq g \leq k - 1,\ s \geq 0,\ v \geq k/10,\ u \geq 0,\ u + v \geq k + 1.$$

Let R be the subset of $\boldsymbol{\xi} \in S_k(u, v)$ where, for some $l \geq g + 1$, we have

$$(4.4) \qquad \frac{l - u}{v} \leq \xi_l \leq \frac{l - u + 1}{v}, \qquad \xi_{l-g} \geq \frac{l - u - s}{v}.$$

Then

$$\operatorname{Vol}(R) \ll \frac{g^2(10(s + 1))^g}{g!}\frac{(u + 1)(u + v - k)^2}{(k + 1)!}.$$

Proof. Fix l satisfying $\max(u, g+1) \leq l \leq k$. Let R_l be the subset of $\boldsymbol{\xi} \in S_k(u, v)$ satisfying (4.4) for this particular l. We have $\mathrm{Vol}(R_l) \leq V_1 V_2 V_3 V_4$, where, by Lemma 4.1,

$$V_1 = \mathrm{Vol}\{0 \leq \xi_1 \leq \cdots \leq \xi_{l-g-1} \leq \tfrac{l-u+1}{v} : \xi_i \geq \tfrac{i-u}{v} \ \forall i\}$$

$$= \left(\frac{l-u+1}{v}\right)^{l-g-1} \mathrm{Vol}\{0 \leq \theta_1 \leq \cdots \leq \theta_{l-g-1} \leq 1 : \theta_i \geq \tfrac{i-u}{l-u+1} \ \forall i\}$$

$$= \left(\frac{l-u+1}{v}\right)^{l-g-1} \frac{Q_{l-g-1}(u, l-u+1)}{(l-g-1)!}$$

$$\ll \left(\frac{l-u+1}{v}\right)^{l-g-1} \frac{(u+1)g^2}{(l-g)!},$$

$$V_2 = \mathrm{Vol}\{\tfrac{l-u-s}{v} \leq \xi_{l-g} \leq \cdots \leq \xi_{l-1} \leq \tfrac{l-u+1}{v}\} = \frac{1}{g!}\left(\frac{s+1}{v}\right)^g,$$

$$V_3 = \mathrm{Vol}\{\tfrac{l-u}{v} \leq \xi_l \leq \tfrac{l-u+1}{v}\} = \frac{1}{v},$$

$$V_4 = \mathrm{Vol}\{\xi_{l+1} \leq \cdots \leq \xi_k \leq 1 : \xi_i \geq \tfrac{i-u}{v} \ \forall i\}$$

$$= \frac{1}{(k-l)!}\left(\frac{u+v-l}{v}\right)^{k-l} Q_{k-l}(0, u+v-l)$$

$$\ll \left(\frac{u+v-l}{v}\right)^{k-l} \frac{(u+v-k)^2}{(k-l+1)!}.$$

Thus

$$\mathrm{Vol}(R) \ll \frac{(s+1)^g(u+1)g^2(u+v-k)^2}{g! v^k (k+1-g)!} \sum_l \binom{k+1-g}{l-g}(l-u+1)^{l-g-1}(u+v-l)^{k-l}.$$

By Lemma 4.2 (with $t = k+1-g$, $a = g+1-u$, $b = u+v-k-1$), the sum on l is

$$\leq e^4(v+1)^{k-g} \ll v^{k-g} = \frac{v^k}{k^g}\left(\frac{k}{v}\right)^g \leq v^k 10^g \frac{(k-g+1)!}{k \cdot k!}$$

and the lemma follows. $\qquad\square$

To bound $U_k(v)$, we will bound the volume of the set

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

Lemma 4.4. *Suppose k, v, γ are integers with $1 \leq k \leq 10v$ and $\gamma \geq 0$. Set $b = k - v$. Then*

$$\mathrm{Vol}(\mathscr{T}(k, v, \gamma)) \ll \frac{Y}{2^{2^{b-\gamma}}(k+1)!}, \qquad Y = \begin{cases} b & \text{if } b \geq \gamma + 5 \\ (\gamma+5-b)^2(\gamma+1) & \text{if } b \leq \gamma + 4 \end{cases}.$$

Proof. Let $r = \max(5, b-\gamma)$ and $\boldsymbol{\xi} \in \mathscr{T}(k, v, \gamma)$. Then either

$$(4.5) \qquad\qquad \xi_j > \tfrac{j-\gamma-r}{v} \quad (1 \leq j \leq k)$$

or

$$(4.6) \qquad \min_{1 \leq j \leq k}(\xi_j - \tfrac{j-\gamma}{v}) = \xi_l - \tfrac{l-\gamma}{v} \in [\tfrac{-h}{v}, \tfrac{1-h}{v}] \text{ for some integers } h \geq r+1, 1 \leq l \leq k.$$

Let V_1 be the volume of $\boldsymbol{\xi} \in \mathscr{T}(k, v, \gamma)$ satisfying (4.5). If $b \geq \gamma + 5$, (4.5) is not possible, so $b \leq \gamma + 4$ and $r = 5$. By Theorem 4.1,

$$V_1 \leq \frac{Q_k(\gamma + 5, v)}{k!} \ll \frac{(\gamma + 6)(\gamma + 6 - b)^2}{(k + 1)!} \ll \frac{Y}{2^{2^{b-\gamma}}(k + 1)!}.$$

If (4.6) holds, then there is an integer m satisfying

(4.7) $m \geq h - 3, \ \xi_{l-2^m} \geq \frac{l - \gamma - 2m}{v}.$

To see (4.7), suppose such an m does not exist. Then

$$2^{v\xi_1} + \cdots + 2^{v\xi_l} \leq 2^{h-3}2^{l-\gamma-h+1} + \sum_{m \geq h-3} 2^m 2^{l-\gamma-2m} \leq 2^{l-\gamma},$$

a contradiction. Let V_2 be the volume of $\boldsymbol{\xi} \in \mathscr{T}(k, v, \gamma)$ satisfying (4.6). Fix h and m satisfying (4.7) and use Lemma 4.3 with $u = \gamma + h$, $g = 2^m$, $s = 2m$. The volume of such $\boldsymbol{\xi}$ is

$$\ll \frac{(\gamma + h + 1)(\gamma + h - b)^2}{(k + 1)!} \frac{(20m + 10)^{2^m} 2^{2m}}{(2^m)!} \ll \frac{(\gamma + h + 1)(\gamma + h - b)^2}{2^{2^{m+3}}(k + 1)!}.$$

The sum of $2^{-2^{m+3}}$ over $m \geq h - 3$ is $\ll 2^{-2^h}$. Summing over $h \geq r + 1$ gives

$$V_2 \ll \frac{(\gamma + r + 2)(\gamma - b + r + 2)^2}{2^{2^{r+1}}(k + 1)!} \ll \frac{Y}{2^{2^{b-\gamma}}(k + 1)!}.$$

$\square$

Proof of Lemma 3.6. Assume $k \geq 1$, since the lemma is trivial when $k = 0$. Put $b = k - v$. For integers $m \geq 0$, consider $\boldsymbol{\xi} \in R_k$ satisfying $2^{-m} \leq \min_{0 \leq j \leq k} 2^{-j} \left(2^{v\xi_1} + \cdots + 2^{v\xi_j} + 1 \right) < 2^{1-m}$. For $1 \leq j \leq k$ we have

$$2^{-j} \left(2^{v\xi_1} + \cdots + 2^{v\xi_j} \right) \geq \max(2^{-j}, 2^{-m} - 2^{-j}) \geq 2^{-m-1},$$

so $\boldsymbol{\xi} \in \mathscr{T}(k, v, m + 1)$. Hence, by Lemma 4.4,

$$U_k(v) \leq \sum_{m \geq 0} 2^{1-m} \operatorname{Vol}(\mathscr{T}(k, v, m + 1)) \ll \frac{1}{(k + 1)!} \sum_{m \geq 0} \frac{2^{-m}Y_m}{2^{2^{b+t-m-1}}},$$

$$Y_m = \begin{cases} b & \text{if } m \leq b - 6 \\ (m + 6 - b)^2(m + 2) & \text{if } m \geq b - 5 \end{cases}.$$

Next,

$$\sum_{m \geq 0} \frac{2^{-m}Y_m}{2^{2^{b-m-1}}} \ll \sum_{0 \leq m \leq b-6} \frac{b}{2^m 2^{2^{b-m-1}}} + \sum_{m \geq \max(0, b-5)} \frac{(m + 6 - b)^2(m + 2)}{2^m}.$$

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

References

1. Sean Eberhard, Kevin Ford, and Ben Green, *Permutations fixing k-sets*, International Mathematical Research Notices **2016** (2016), no. 21, 6713–6731.
2. P. Erdős, *Some remarks on number theory*, Riveon Lematematika **9** (1955), 45–48, (Hebrew. English summary). MR 17,460d
3. _____, *An asymptotic inequality in the theory of numbers*, Vestnik Leningrad. Univ. **15** (1960), no. 13, 41–49, (Russian). MR 23 #A3720
4. K. Ford, *The distribution of integers with a divisor in a given interval*, Ann. Math. (2008), 367–433.
5. H. Halberstam and H.-E. Richert, *Sieve methods*, Academic Press [A subsidiary of Harcourt Brace Jovanovich, Publishers], London-New York, 1974, London Mathematical Society Monographs, No. 4. MR 54 #12689
6. R. R. Hall and G. Tenenbaum, *Divisors*, Cambridge Tracts in Mathematics, vol. 90, Cambridge University Press, Cambridge, 1988. MR 90a:11107
7. Dimitrios Koukoulopoulos, *Localized factorizations of integers*, Proc. London Math. Soc. **101** (2010), 392–426.
8. J. Riordan, *Combinatorial identities*, John Wiley & Sons Inc., New York, 1968. MR 38 #53

DEPARTMENT OF MATHEMATICS, 1409 WEST GREEN STREET, UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN, URBANA, IL 61801, USA

E-mail address: ford@math.uiuc.edu