Annals of Mathematics, **168** (2008), 367–433

The distribution of integers with a divisor in a given interval

By Kevin Ford

Abstract

We determine the order of magnitude of $H(x, y, z)$, the number of integers $n \le x$ having a divisor in $(y, z]$, for all x, y and z. We also study $H_r(x, y, z)$, the number of integers $n \le x$ having exactly r divisors in $(y, z]$. When $r = 1$ we establish the order of magnitude of $H_1(x, y, z)$ for all x, y, z satisfying $z \le x^{1/2-\varepsilon}$. For every $r \ge 2$, $C > 1$ and $\varepsilon > 0$, we determine the order of magnitude of $H_r(x, y, z)$ uniformly for y large and $y + y/(\log y)^{\log 4-1-\varepsilon} \le z \le \min(y^C, x^{1/2-\varepsilon})$. As a consequence of these bounds, we settle a 1960 conjecture of Erdős and some conjectures of Tenenbaum. One key element of the proofs is a new result on the distribution of uniform order statistics.

Contents

1. Introduction

For $0 < y < z$, let $\tau(n; y, z)$ be the number of divisors d of n which satisfy $y < d \le z$. Our focus in this paper is to estimate $H(x, y, z)$, the number of positive integers $n \le x$ with $\tau(n; y, z) > 0$, and $H_r(x, y, z)$, the number of

$n \le x$ with $\tau(n; y, z) = r$. By inclusion-exclusion,

$$H(x,y,z) = \sum_{k\ge 1} (-1)^{k-1} \sum_{y<d_1<\cdots<d_k\le z} \left\lfloor \frac{x}{\operatorname{lcm}[d_1,\cdots,d_k]} \right\rfloor,$$

but this is not useful for estimating $H(x,y,z)$ unless $z-y$ is small. With y and z fixed, however, this formula implies that the set of positive integers having at least one divisor in $(y,z]$ has an *asymptotic density*, i.e. the limit

$$\varepsilon(y,z) = \lim_{x\to\infty} \frac{H(x,y,z)}{x}$$

exists. Similarly, the exact formula

$$H_r(x,y,z) = \sum_{k\ge r} (-1)^{k-r} \binom{k}{r} \sum_{y<d_1<\cdots<d_k\le z} \left\lfloor \frac{x}{\operatorname{lcm}[d_1,\cdots,d_k]} \right\rfloor$$

implies the existence of

$$\varepsilon_r(y,z) = \lim_{x\to\infty} \frac{H_r(x,y,z)}{x}$$

for every fixed pair y, z.

1.1. *Bounds for $H(x,y,z)$.* Besicovitch initiated the study of such quantities in 1934, proving in [2] that

$$\liminf_{y\to\infty} \varepsilon(y,2y) = 0, \tag{1.1}$$

and using (1.1) to construct an infinite set $\mathscr{A}$ of positive integers such that its set of multiples $\mathscr{B}(\mathscr{A}) = \{am : a \in \mathscr{A}, m \ge 1\}$ does not possess asymptotic density. Erdős in 1935 [5] showed $\lim_{y\to\infty} \varepsilon(y,2y) = 0$ and in 1960 [8] gave the further refinement (see also Tenenbaum [38])

$$\varepsilon(y,2y) = (\log y)^{-\delta+o(1)} \qquad (y\to\infty),$$

where

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

Prior to the 1980s, a few other special cases were studied. In 1936, Erdős [6] established

$$\lim_{y\to\infty} \varepsilon(y, y^{1+u}) = 0,$$

provided that $u = u(y) \to 0$ as $y\to\infty$. In the late 1970s, Tenenbaum ([39], [40]) showed that

$$h(u,t) = \lim_{x\to\infty} \frac{H(x, x^{(1-u)/t}, x^{1/t})}{x}$$

exists for $0 \le u \le 1$, $t \ge 1$ and gave bounds on $h(u,t)$.

Motivated by a growing collection of applications for such bounds, Tenenbaum in the early 1980s turned to the problem of bounding $H(x, y, z)$ for *all* x, y, z. In the seminal work [42] he established reasonably sharp upper and lower bounds for $H(x, y, z)$ which we list below (paper [41] announces these results and gives a history of previous bounds for $H(x, y, z)$; Hall and Tenenbaum's book *Divisors* [24] gives a simpler proof of Tenenbaum's theorem). We require some additional notation. For a given pair (y, z) with $4 \le y < z$, we define η, u, β, ξ by

$$z = e^\eta y = y^{1+u}, \quad \eta = (\log y)^{-\beta}, \quad \beta = \log 4 - 1 + \frac{\xi}{\sqrt{\log\log y}}. \tag{1.2}$$

Tenenbaum defines η by $z = y(1 + \eta)$, which is asymptotic to our η when $z - y = o(y)$. The definition in (1.2) plays a natural role in the arguments even when $z - y$ is large. For smaller z, we also need the function

$$G(\beta) = \begin{cases} \frac{1+\beta}{\log 2} \log\left(\frac{1+\beta}{e\log 2}\right) + 1 & 0 \le \beta \le \log 4 - 1 \\ \beta & \log 4 - 1 \le \beta. \end{cases} \tag{1.3}$$

When x and y are fixed, Tenenbaum discovered that $H(x, y, z)$ undergoes a change of behavior in the vicinity of

$$z = z_0(y) := y \exp\{(\log y)^{1-\log 4}\} \approx y + y/(\log y)^{\log 4 - 1},$$

in the vicinity of $z = 2y$ and in the vicinity of $z = y^2$.

Theorem T1 (Tenenbaum [42]). (i) *Suppose* $y \to \infty$, $z - y \to \infty$, $z \le \sqrt{x}$ *and* $\xi \to \infty$. *Then*

$$H(x, y, z) \sim \eta x.$$

(ii) *Suppose* $2 \le y < z \le \min(2y, \sqrt{x})$ *and* ξ *is bounded above. Then*

$$\frac{x}{(\log y)^{G(\beta)} Z(\log y)} \ll H(x, y, z) \ll \frac{x}{(\log y)^{G(\beta)} \max(1, -\xi)}.$$

Here $Z(v) = \exp\{c\sqrt{\log(100v)\log\log(100v)}\}$ *and* c *is some positive constant.*

(iii) *Suppose* $4 \le 2y \le z \le \min(y^{3/2}, \sqrt{x})$. *Then*

$$\frac{xu^\delta}{Z(1/u)} \ll H(x, y, z) \ll \frac{xu^\delta \log\log(3/u)}{\sqrt{\log(2/u)}}.$$

Moreover, the term $\log\log(3/u)$ *on the right may be omitted if* $z \le By$ *for some* $B > 2$, *the constant implied by* $\ll$ *depending on* B.

(iv) *If* $2 \le y \le z \le x$, *then*

$$H(x, y, z) = x\left(1 + O\left(\frac{\log y}{\log z}\right)\right).$$

Remark. Since

$$\sum_{n\le x}\tau(n,y,z)=\sum_{y<d\le z}\left\lfloor\frac{x}{d}\right\rfloor\sim\eta x\qquad(z-y\to\infty),$$

in the range of x,y,z given in (i) of Theorem T1, most n with a divisor in $(y,z]$ have only one such divisor. By (iv), when $\frac{\log z}{\log y}\to\infty$, almost all integers have a divisor in $(y,z]$.

In 1991, Hall and Tenenbaum [25] established the order of $H(x,y,z)$ in the vicinity of the "threshhold" $z=z_0(y)$. Specifically, they showed that if $3\le y+1\le z\le\sqrt{x}$, $c>0$ is fixed and $\xi\ge -c(\log\log y)^{1/6}$, then

$$H(x,y,z)\asymp\frac{x}{(\log y)^{G(\beta)}\max(1,-\xi)},$$

thus showing that the upper bound given by (ii) of Theorem T1 is the true order in this range. In fact the argument in [25] implies that

$$H_1(x,y,z)\asymp H(x,y,z)$$

in this range of x,y,z. Specifically, Hall and Tenenbaum use a lower estimate

$$H(x,y,z)\ge\sum_{\substack{n\le x\\ n\in\mathscr{N}}}\tau(n,y,z)(2-\tau(n,y,z))$$

for a certain set $\mathscr{N}$, and clearly the right side is also a lower bound for $H_1(x,y,z)$. Later, in a slightly more restricted range, Hall ([22], Ch. 7) proved an asymptotic formula for $H(x,y,z)$ which extends the asymptotic formula of part (i) of Theorem T1. Richard Hall has kindly pointed out an error in the stated range of validity of this asymptotic in [22], which we correct below (in [22], the range is stated as $\xi\ge -c(\log\log y)^{1/6}$).

Theorem H (Hall [22, Th. 7.9]). *Uniformly for $z\le x^{1/\log\log x}$ and for $\xi\ge -o(\log\log y)^{1/6}$,*

$$\frac{H(x,y,z)}{x}=(F(\xi)+O(E(\xi,y)))(\log y)^{-\beta},$$

where

$$F(\xi)=\frac{1}{\sqrt{\pi}}\int_{-\infty}^{\xi/\log 4}e^{-u^2}\,du$$

and

$$E(\xi,y)=\begin{cases}\dfrac{\xi^2+\log\log\log y}{\sqrt{\log\log y}}e^{-\xi^2/\log^2 4}, & \xi\le 0\\[2ex] \dfrac{\xi+\log\log\log y}{\sqrt{\log\log y}}, & \xi>0.\end{cases}$$

Note that
$$F(\xi)(\log y)^{-\beta} \asymp \frac{1}{(\log y)^{G(\beta)} \max(1,-\xi)}$$
in Theorem H.

We now determine the exact order of $H(x,y,z)$ for all x,y,z. Constants implied by O, $\ll$ and $\asymp$ are absolute unless otherwise noted, e.g. by a subscript. The notation $f \asymp g$ means $f \ll g$ and $g \ll f$. Variables $c_1, c_2, \dots$ will denote certain specific constants, y_0 is a sufficiently large real number, while $y_0(\cdot)$ will denote a large constant depending only on the parameters given, e.g. $y_0(r,c,c')$, and the meaning may change from statement to statement. Lastly, $\lfloor x \rfloor$ denotes the largest integer $\le x$.

THEOREM 1. *Suppose* $1 \le y \le z \le x$. *Then*,

(i) $H(x,y,z) = 0$ *if* $z < \lfloor y \rfloor + 1$;

(ii) $H(x,y,z) = \lfloor x/(\lfloor y \rfloor + 1) \rfloor$ *if* $\lfloor y \rfloor + 1 \le z < y+1$;

(iii) $H(x,y,z) \asymp 1$ *if* $z \ge y+1$ *and* $x \le 100000$;

(iv) $H(x,y,z) \asymp x$ *if* $x \ge 100000$, $1 \le y \le 100$ *and* $z \ge y+1$;

(v) *If* $x > 100000$, $100 \le y \le z-1$ *and* $y \le \sqrt{x}$,
$$\frac{H(x,y,z)}{x} \asymp \begin{cases} \log(z/y) = \eta & y+1 \le z \le z_0(y) \\ \dfrac{\beta}{\max(1,-\xi)(\log y)^{G(\beta)}} & z_0(y) \le z \le 2y \\ u^{\delta}(\log \frac{2}{u})^{-3/2} & 2y \le z \le y^2 \\ 1 & z \ge y^2. \end{cases}$$

(vi) *If* $x > 100000$, $\sqrt{x} < y < z \le x$ *and* $z \ge y+1$, *then*
$$H(x,y,z) \asymp \begin{cases} H\left(x, \frac{x}{z}, \frac{x}{y}\right) & \frac{x}{y} \ge \frac{x}{z} + 1 \\ \eta x & \textit{otherwise.} \end{cases}$$

COROLLARY 1. *Suppose* $x_1, y_1, z_1, x_2, y_2, z_2$ *are real numbers with* $1 \le y_i < z_i \le x_i$ $(i=1,2)$, $z_i \ge y_i + 1$ $(i=1,2)$, $\log(z_1/y_1) \asymp \log(z_2/y_2)$, $\log y_1 \asymp \log y_2$ *and* $\log(x_1/z_1) \asymp \log(x_2/z_2)$. *Then*
$$\frac{H(x_1,y_1,z_1)}{x_1} \asymp \frac{H(x_2,y_2,z_2)}{x_2}.$$

Corollary 2. *If $c > 1$ and $\frac{1}{c-1} \le y \le x/c$, then*

$$H(x, y, cy) \asymp_c \frac{x}{(\log Y)^{\delta}(\log\log Y)^{3/2}} \qquad (Y = \min(y, x/y) + 3)$$

and

$$\varepsilon(y, cy) \asymp_c \frac{1}{(\log y)^{\delta}(\log\log y)^{3/2}}.$$

Items (i)–(iv) of Theorem 1 are trivial. The first and fourth part of item (v) are already known (cf. the papers of Tenenbaum [42] and Hall and Tenenbaum [25] mentioned above). Item (vi) essentially follows from (v) by observing that $d|n$ if and only if $(n/d)|n$. However, proving (vi) requires a version of (v) where n is restricted to a short interval, which we record below. The range of Δ can be considerably improved, but the given range suffices for the application to Theorem 1 (vi).

Theorem 2. *For $y_0 \le y \le \sqrt{x}$, $z \ge y + 1$ and $\frac{x}{\log^{10} z} \le \Delta \le x$,*

$$H(x, y, z) - H(x - \Delta, y, z) \asymp \frac{\Delta}{x} H(x, y, z).$$

Motivated by an application to gaps in the Farey series, we also record an analogous result for $H^*(x, y, z)$, the number of *squarefree* numbers $n \le x$ with $\tau(n, y, z) \ge 1$.

Theorem 3. *Suppose $y_0 \le y \le \sqrt{x}$, $y + 1 \le z \le x$ and $\frac{x}{\log y} \le \Delta \le x$. If $z \ge y + Ky^{1/5} \log y$, where K is a large absolute constant, then*

$$H^*(x, y, z) - H^*(x - \Delta, y, z) \asymp \frac{\Delta}{x} H(x, y, z).$$

If $y + (\log y)^{2/3} \le z \le y + Ky^{1/5} \log y$, $g > 0$ and there are $\ge g(z-y)$ square-free numbers in $(y, z]$, then

$$H^*(x, y, z) - H^*(x - \Delta, y, z) \asymp_g \frac{\Delta}{x} H(x, y, z).$$

To obtain good lower bounds on $H^*(x, y, z)$, it is important that $(y, z]$ contain many squarefree integers. In the extreme case where $(y, z]$ contains no squarefree integers, clearly $H^*(x, y, z) = 0$. A theorem of Filaseta and Trifonov [13] implies that there are $\ge \frac{1}{2}(z - y)$ squarefree numbers in $(y, z]$ if $z \ge y + Ky^{1/5} \log y$, and this is the best result known of this kind.

Some applications. Most of the following applications depend on the distribution of integers with $\tau(n, y, z) \ge 1$ when $z \asymp y$. See also Chapter 2 of [24] for further discussion of these and other applications.

1. Distinct products in a multiplication table, a problem of Erdős from 1955 ([7], [8]). Let $A(x)$ be the number of positive integers $n \le x$ which can be written as $n = m_1 m_2$ with each $m_i \le \sqrt{x}$.

Corollary 3. *We have*

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

Proof. Apply Theorem 1 and the inequalities

$$H\left(\frac{x}{4}, \frac{\sqrt{x}}{4}, \frac{\sqrt{x}}{2}\right) \le A(x) \le \sum_{k\ge 0} H\left(\frac{x}{2^k}, \frac{\sqrt{x}}{2^{k+1}}, \frac{\sqrt{x}}{2^k}\right). \qquad \square$$

2. Distribution of Farey gaps (Cobeli, Ford, Zaharescu [3]).

Corollary 4. *Let* $(\frac{0}{1}, \frac{1}{Q}, \dots, \frac{Q-1}{Q}, \frac{1}{1})$ *denote the sequence of Farey fractions of order* Q, *and let* $N(Q)$ *denote the number of distinct gaps between successive terms of the sequence. Then*

$$N(Q) \asymp \frac{Q^2}{(\log Q)^{\delta}(\log\log Q)^{3/2}}.$$

Proof. The distinct gaps are precisely those products qq' with $1 \le q$, $q' \le Q$, $(q, q') = 1$ and $q + q' > Q$. Thus

$$H^*(\tfrac{9}{25}Q^2, \tfrac{Q}{2}, \tfrac{3Q}{5}) - H^*(\tfrac{3}{10}Q^2, \tfrac{Q}{2}, \tfrac{3Q}{5}) \le N(Q) \le H(Q^2, Q/2, Q),$$

and the corollary follows from Theorems 1 and 3. $\square$

3. Divisor functions. Erdős introduced ([11], [12] and §4.6 of [24]) the function

$$\tau^+(n) = |\{k \in \mathbb{Z} : \tau(n, 2^k, 2^{k+1}) \ge 1\}|.$$

Corollary 5. *For* $x \ge 3$,

$$\frac{1}{x}\sum_{n\le x} \tau^+(n) \asymp \frac{(\log x)^{1-\delta}}{(\log\log x)^{3/2}}.$$

Proof. This follows directly from Theorem 1 and

$$\sum_{n\le x} \tau^+(n) = \sum_k H(x, 2^k, 2^{k+1}). \qquad \square$$

Tenenbaum [37] defines $\rho_1(n)$ to be the largest divisor d of n satisfying $d \le \sqrt{n}$.

Corollary 6. *We have*

$$\sum_{n\le x} \rho_1(n) \asymp \frac{x^{3/2}}{(\log x)^{\delta}(\log\log x)^{3/2}}.$$

Proof. Suppose $x/4^l < n \le x/4^{l-1}$. Since $\rho_1(n)$ lies in $(\sqrt{x}2^{-k}, \sqrt{x}2^{-k+1}]$ for some integer $k \ge l$,

$$\frac{\sqrt{x}}{4}\left(H\left(x, \tfrac{\sqrt{x}}{4}, \tfrac{\sqrt{x}}{2}\right) - H\left(\tfrac{x}{4}, \tfrac{\sqrt{x}}{4}, \tfrac{\sqrt{x}}{2}\right)\right) \le \sum_{n\le x} \rho_1(n)$$
$$\le \sum_{l\ge 1}\sum_{k\ge l} \frac{\sqrt{x}}{2^{k-1}} H\left(\tfrac{x}{4^{l-1}}, \tfrac{\sqrt{x}}{2^k}, \tfrac{\sqrt{x}}{2^{k-1}}\right)$$

and the corollary follows from Theorem 1. □

4. Density of unions of residue classes. Given moduli $m_1, \dots, m_k$, let $\delta_0(m_1, \dots, m_k)$ be the minimum, over all possible residue classes $a_1 \bmod m_1$, $\dots, a_k \bmod m_k$, of the density of integers which lie in at least one of the classes. By a theorem of Rogers (see [20, p. 242–244]), the minimum is achieved by taking $a_1 = \cdots = a_k = 0$ and thus $\delta_0(m_1, \dots, m_k)$ is the density of integers possessing a divisor among the numbers $m_1, \dots, m_k$. When $m_1, \dots, m_k$ consist of the integers in an interval $(y, z]$, then $\delta_0(m_1, \dots, m_k) = \varepsilon(y, z)$.

5. Bounds for $H(x, y, z)$ were used in recent work of Heath-Brown [26] on the validity of the Hasse principle for pairs of quadratic forms.

6. Bounds on $H(x, y, z)$ are central to the study of the function

$$\max\{|a-b| : 1 \le a, b \le n-1, ab \equiv 1 \pmod{n}\}$$

in [16].

1.2. *Bounds for $H_r(x, y, z)$.* In the paper [8], Erdős made the following conjecture:[1]

CONJECTURE 1 (Erdős [8]).

$$\lim_{y\to\infty} \frac{\varepsilon_1(y, 2y)}{\varepsilon(y, 2y)} = 0.$$

This can be interpreted as the assertion that the conditional probability that a random integer has exactly 1 divisor in $(y, 2y]$ given that it has at least one divisor in $(y, 2y]$, tends to zero as $y \to \infty$.

In 1987, Tenenbaum [43] gave general bounds on $H_r(x, y, z)$, which are of similar strength to his bounds on $H(x, y, z)$ (Theorem T1) when $z \le 2y$.

THEOREM T2 (Tenenbaum [43]). *Fix $r \ge 1$, $c > 0$.*

[1] Erdős also mentioned this conjecture in some of his books on unsolved problems, e.g. [9], and he wrote it in the Problem Book (page 2) of the Mathematisches Forschungsinstitut Oberwolfach.

(i) *If* $y \to \infty$, $z - y \to \infty$, *and* $\xi \to \infty$, *then*

$$\frac{H_r(x,y,z)}{H(x,y,z)} \to \begin{cases} 1 & r = 1 \\ 0 & r \ge 2 \end{cases}.$$

(ii) *If* $y \ge y_0(r)$, $z_0(y) \le z \le \min(2y, x^{1/(r+1)-c})$, *then*

$$\frac{1}{Z(\log y)} \ll_{r,c} \frac{H_r(x,y,z)}{H(x,y,z)} \le 1.$$

(iii) *If* $y_0(r) \le 2y \le z \le \min(y^{3/2}, x^{1/(r+1)-c})$,

$$\frac{1}{\log(z/y) Z(\log y)} \ll_{r,c} \frac{H_r(x,y,z)}{H(x,y,z)} \ll_r \frac{Z(\log y)}{(\log(z/y))^{\delta}}.$$

(iv) *If* $y \ge y_0(r)$, $y^{3/2} \le z \le x^{1/2}$, *then*

$$\frac{\log\left(\frac{\log z}{\log y}\right)}{\log z} \ll_r \frac{H_r(x,y,z)}{H(x,y,z)} \ll_r \frac{(\log y)^{1-\delta}(\log\log z)^{2r+1}}{\log z}.$$

Remarks. In [43], (ii) and (iii) above are stated with $c = 0$, but the proofs of the lower bounds require c to be positive. The construction of n with $\tau(n,y,z) = r$ on p. 177 of [43] requires $z^{\frac{1}{r+3}+r+1} \le x$, but the proof can be modified to work for $z \le x^{\frac{1}{r+1}-c}$ for any fixed $c > 0$.

Based on the strength of the bounds in (ii) and (iii) above, Tenenbaum made two conjectures. In particular, he asserted that Conjecture 1 is false.

CONJECTURE 2 (Tenenbaum [43]). *For every* $r \ge 1$, $c > 0$, *and* $c' > 0$, *if* $\xi \to -\infty$ *as* $y \to \infty$, $y \le x^{1/2-c'}$ *and* $z \le cy$, *then*

$$H_r(x,y,z) \gg_{r,c,c'} H(x,y,z).$$

CONJECTURE 3 (Tenenbaum [43]). *If* $c > 0$ *is fixed,* $y \le x^{1/2-c}$, $r \ge 1$ *and* $z/y \to \infty$, *then*

$$H_r(x,y,z) = o(H(x,y,z)).$$

Using the methods used to prove Theorem 1 plus some additional arguments, we shall prove much stronger bounds on $H_r(x,y,z)$ which will settle these three conjectures (except Conjecture 2 when z is near $z_0(y)$). When $z \ge 2y$, the order of $H_r(x,y,z)$ depends on $\nu(r)$, the exponent of the largest power of 2 dividing r (i.e. $2^{\nu(r)} \| r$).

THEOREM 4. *Suppose that* $c > 0$, $y_0(c) \le y$, $y + 1 \le z \le x^{5/8}$ *and* $yz \le x^{1-c}$. *Then*

$$\frac{H_1(x,y,z)}{H(x,y,z)} \asymp_c \frac{\log\log(z/y+10)}{\log(z/y+10)}. \tag{1.4}$$

Theorem 5. *Suppose that* $r \ge 2$, $c > 0$, $y_0(r,c) \le y$, $z \le x^{5/8}$ *and* $yz \le x^{1-c}$. *If* $z_0(y) \le z \le 10y$, *then*

$$\frac{\max(1,-\xi)}{\sqrt{\log\log y}} \ll_{r,c} \frac{H_r(x,y,z)}{H(x,y,z)} \le 1. \tag{1.5}$$

When $C > 1$ *is fixed and* $10y \le z \le y^C$,

$$\frac{H_r(x,y,z)}{H(x,y,z)} \asymp_{r,c,C} \frac{(\log\log(z/y))^{\nu(r)+1}}{\log(z/y)}. \tag{1.6}$$

When $y \ge y_0(r)$ *and* $y^2 \le z \le x^{5/8}$, *then*

$$\frac{H_r(x,y,z)}{H(x,y,z)} \gg_r \frac{(\log\log y)^{\nu(r)+1}}{\log z}. \tag{1.7}$$

Corollary 7. *For every* $\lambda > 1$ *and* $r \ge 1$,

$$\frac{\varepsilon_r(y,\lambda y)}{\varepsilon(y,\lambda y)} \gg_{r,\lambda} 1.$$

while for each $r \ge 1$, *if* $z/y \to \infty$ *then*

$$\frac{\varepsilon_r(y,z)}{\varepsilon(y,z)} \to 0.$$

In particular, Conjecture 1 is false, Conjecture 3 is true, and Conjecture 2 is true provided $z \ge y + y/(\log y)^{\log 4 - 1 - b}$ for a fixed $b > 0$.

The upper bounds in Theorems 4 and 5 are proved in the wider range $y \le \sqrt{x}, z \le x^{5/8}$. The conclusions of the two theorems, however, are not true when $yz \approx x$. This is a consequence of $d|n$ implying $\frac{n}{d}|n$, which shows for example that $\tau(n,y,n/y)$ is odd only if n is a square or $y|n$. For another example, while $H_1(x,x^{1/4},x^{3/5}) \asymp x\frac{\log\log x}{\log x}$ by Theorem 4, we have

$$H_1(x,x^{1/4},x^{3/4}) \asymp \frac{x}{\log x}. \tag{1.8}$$

The lower bound is obtained by considering $n = ap$ with $a \le x^{1/4}$ and p a prime in $(\frac12 x^{3/4}, x^{3/4}]$. Now suppose $n > x^{3/4}$, $\tau(n,x^{1/4},x^{3/4}) = 1$, and $d|n$ with $x^{1/4} < d \le x^{3/4}$. Since $\frac{n}{d} < x^{3/4}$, we have $\frac{n}{d} \le x^{1/4}$, hence $d > x^{1/2}$. If d is not prime, then there is a proper divisor of d that is $\ge \sqrt{d} > x^{1/4}$, a contradiction. Thus, d is prime and $n = da$ with $a \le x^{1/4}$. The upper bound in (1.8) follows.

There is an application to the Erdős-Montgomery function $g(n)$, which counts the number of pairs of consecutive divisors d, d' of n with $d|d'$ (see [11], [12]). The following sharpens Théorème 2 of [43].

COROLLARY 8. *We have*

$$\frac{1}{x}\sum_{n\le x} g(n) \asymp \frac{(\log x)^{1-\delta}}{(\log\log x)^{3/2}}.$$

Proof. The upper bound follows from $g(n) \le \tau^+(n)$ and Corollary 5. We also have $g(2n) \ge I(n)$, where $I(n)$ is the number of $d|n$ such that if $d'|n, d' \ne d$, then $d' > 2d$ or $d' < d/2$ (see §4). We quickly derive (cf. [43, p. 185])

$$\sum_{n\le x} g(n) \gg \frac{x}{\log x}\sum_{m\le x^{1/5}} \frac{I(m)}{m}.$$

Applying (5.5) with $g = 1$, $y = \sqrt{x}$, $\alpha = \frac{1}{5}$ and $\sigma = \log 2$, we obtain

$$\sum_{m\le x^{1/5}} \frac{I(m)}{m} \gg \frac{(\log x)^{2-\delta}}{(\log\log x)^{3/2}},$$

and this gives the corollary. □

In a forthcoming paper, the author and G. Tenenbaum [18] show that Conjecture 2 is false when z is close to $z_0(y)$. Specifically, if $c > 0$ is fixed, $g(y) > 0$, $\lim_{y\to\infty} g(y) = 0$, $y \le x^{1/2-c}$ and $y + 1 \le z \le y + y(\log y)^{1-\log 4+g(y)}$, then

$$H_1(x, y, z) \sim H(x, y, z) \qquad (y \to \infty).$$

Moreover, the lower bound in (1.5) is the true order of $\frac{H_r(x,y,z)}{H(x,y,z)}$ for $r \ge 2$.

1.3. *Divisors of shifted primes.* The methods developed in this paper may also be used to estimate a more general quantity

$$H(x, y, z; \mathscr{A}) = |\{n \le x : n \in \mathscr{A}, \tau(n, y, z) \ge 1\}$$

for a set $\mathscr{A}$ of positive integers which is well enough distributed in arithmetic progressions so that the initial reductions (Lemmas 3.2, 4.1, 4.2) can be made to work. An example is $\mathscr{A}$ being an arithmetic progression $\{un + v : n \ge 1\}$, where the modulus u may be fixed or grow at a moderate rate as a function of x. Estimates with these $\mathscr{A}$ are given in [16].

One example which we shall examine in this paper is when $\mathscr{A}$ is a set of *shifted primes* (the set $P_\lambda = \{q + \lambda : q \text{ prime}\}$ for a fixed non-zero λ). Results about the multiplicative structure of shifted primes play an important role in many number theoretic applications, especially in the areas of primality testing, integer factorization and cryptography. Upper bounds for $H(x, y, z; P_\lambda)$ have been given by Pappalardi ([32, Th. 3.1]), Erdős and Murty ([10, Th. 2]) and Indlekofer and Timofeev ([27, Th. 2 and its corollaries]). Improving on all of these estimates, we give upper bounds of the expected order of magnitude, for all x, y, z satisfying $y \le \sqrt{x}$.

Theorem 6. *Let λ be a fixed non-zero integer. Let $1 \le y \le \sqrt{x}$ and $y+1 \le z \le x$. Then*

$$H(x,y,z;P_\lambda) \ll_\lambda \begin{cases} \dfrac{H(x,y,z)}{\log x} & z \ge y + (\log y)^{2/3} \\ \\ \dfrac{x}{\log x} \sum_{y<d\le z} \dfrac{1}{\phi(d)} & y < z \le y + (\log y)^{2/3}. \end{cases}$$

Lower bounds are much more difficult, depending heavily on the distribution of primes in arithmetic progressions. The special case $z = y+1$ already presents difficulties, since then $H(x,y,y+1;P_\lambda)$ counts the primes $\le x$ in the progression $-\lambda \pmod{\lfloor y \rfloor + 1}$. If the interval $(y,z]$ is long, however, we can make use of average result for primes in arithmetic progressions.

Theorem 7. *For fixed λ, a, b with $\lambda \ne 0$ and $0 \le a < b \le 1$,*

$$H(x,x^a,x^b;P_\lambda) \gg_{a,b,\lambda} \frac{x}{\log x}.$$

Theorem 7 has an application to counting finite fields for which there is a curve with Jacobian of small exponent [17].

1.4. *Outline of the paper.* In Section 2 we give a few preliminary lemmas about primes and sieve counting functions. Sections 3 and 4 provide an outline of the upper and lower bound arguments with most proofs omitted. These tools are combined to prove Theorems 1, 2, 3, 4 and 5 in Section 5.

The first step in all estimations is to relate the average behavior of $\tau(n,y,z)$, which contains local information about the divisors of n, with average behavior of functions which measure global distribution of divisors. This is accomplished in Section 6. The upper and lower bound arguments begin to diverge after this point. In general, the upper bounds are more difficult, since one may restrict ones attention to numbers with nice properties for the lower bounds. The prime divisors of n which are $< z/y$ play an insignificant role in the estimation of $H(x,y,z)$. For example, if $y < d \le 2y$, then $md \le z$ for $1 \le m \le z/(2y)$. By the same reason, the prime factors of n which are $\le z/y$ play a very important role in the estimation of $H_r(x,y,z)$. Quantifying this difference of roles for the upper bounds in Section 7 is much more difficult than for the lower bounds in Section 9, although the underlying idea is the same.

In Section 7, both $H(x,y,z)$ and $H_r(x,y,z)$ are bounded above in terms of a quantity $S^*(t;\eta)$, which is an average over square-free n whose prime factors lie in $(z/y,z]$ of a global divisor function of n. The contribution to $S^*(t;\eta)$ from those n with exactly k prime factors is then estimated in terms of an integral over $\mathbb{R}^k$ of an elementary but complicated function. Strong estimates for this integral are proved in Section 13, and depend on new probability bounds for uniform order statistics given in Lemma 11.1 (see §11 for relevant definitions).

The lower bound argument follows roughly the same outline as the upper bound, but the details are quite different. Averages over the 'global' divisor functions are estimated in terms of averages of a function which counts 'isolated' divisors of numbers (divisors which are not too close to other divisors) in Section 9. Averages over counts of isolated divisors of numbers with k prime factors are bounded below in terms of the volume of a certain complicated region in $\mathbb{R}^k$. Bounding from below the volume of this region makes use of the bounds on uniform order statistics from Section 11, and this is accomplished in Section 12. For $z \ge y + y(\log y)^{1-\log 4+b}$, $b > 0$ fixed, we need only take a single value of k.

There is an alternative approach to obtaining lower bounds for $H(x, y, z)$ which avoids the use of bounds for order statistics (see §2 of [14]), but they appear to be necessary for our upper bounds and for our lower bounds for $H_r(x, y, z)$.

Finally in Section 14, we apply the upper bound tools developed in the prior sections to give upper estimates for $H(x, y, z; P_\lambda)$, proving Theorem 6. Theorem 7 is much simpler and has a self-contained proof in Section 14.

A relatively short, self-contained proof that

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

is given in [14]. Aside from part of the lower bound argument, the methods are those given here, omitting complications which arise in the general case.

1.5. *Heuristic arguments for $H(x, y, z)$.* Since the prime factors of n which are $< z/y$ play a very insignificant role, we essentially must count how many $n \le x$ have $\tau(n', y, z) \ge 1$, where n' is the product of the primes dividing n lying between z/y and z. For simplicity, assume n' is squarefree, $n' \le z^{100}$ and has k prime factors. When $z \ge y + y(\log y)^{1-\log 4+c}$, $c > 0$ fixed, the majority of such n satisfy $k - k_0 = O(1)$, where

$$k_0 = \left\lfloor \frac{\log\log z - \log\eta}{\log 2} \right\rfloor.$$

For example, most integers n which have a divisor in $(y, 2y]$ have $\frac{\log\log y}{\log 2} + O(1)$ prime factors $\le 2y$.

To see this, assume for the moment that the set $D(n') = \{\log d : d|n'\}$ is uniformly distributed in $[0, \log n_1]$. Then the probability that $\tau(n', y, z) \ge 1$ should be about $2^k \frac{\eta}{\log n_1} \asymp 2^k \frac{\eta}{\log z}$. This is $\gg 1$ precisely when $k \ge k_0 + O(1)$. Using the fact (e.g. Theorem 08 of [24]) that the number of $n \le x$ with n' having k prime factors is approximately

$$\frac{x \log(z/y+2)}{\log z} \frac{(\log\log z - \log\log(z/y+2))^k}{k!},$$

we obtain a heuristic estimate for $H(x, y, z)$ which matches the upper bounds of Theorem T1, sans the $\log\log(3/u)$ factor in (iii). When $\beta = o(1)$ or $\eta > 1$, this is slightly too big. 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 close divisors and large gaps between them. This substantially decreases the likelihood that $\tau(n', y, z) \ge 1$. The cause is slight irregularities in the distribution of prime factors of n' which are guaranteed "almost surely" by large deviation results of probability theory (see e.g. Ch. 1 of [24]). The numbers $\log\log p$ over $p|n'$ are well-known to behave like random numbers in $[\log\log\max(2, z/y), \log\log z]$, and any prime that is slightly below its expectation leads to "clumpiness" in $D(n')$. What we really should count is the number of n for which n' has k prime factors *and* $D(n')$ is roughly uniformly distributed. This corresponds to asking for the prime divisors of n' to lie all above their expected values. An analogy from probability theory is to ask for the likelihood that a random walk on the real numbers, with each step haveing zero expectation, stays completely to the right of the origin (or a point just to the the left of the origin) after k steps. In Section 11 we give estimates for this probability. In the case $z = 2y$, the desired probability is about $1/k \asymp 1/\log\log y$, which accounts for the discrepancy between the upper estimates in Theorem T1 and the bounds in Theorem 1.

1.6. *Some open problems.* (i) Strengthen Theorem 1 to an asymptotic formula.

(ii) Determine the order of $H_r(x, y, z)$ for $r \ge 2$ and $z \ge y^C$ (see the conjecture at the end of §5).

(iii) Establish the order of $H_r(x, y, z)$ when $yz \ge x^{1-c}$.

(iv) Make the dependence on r explicit in Theorem 5 and Corollary 7. Hall and Tenenbaum ([24], Ch. 2) conjecture that for each $r \ge 2$,

$$\lim_{y\to\infty} \frac{\varepsilon_r(y, 2y)}{\varepsilon(y, 2y)} = d_r > 0.$$

In light of (1.6), the sequence $d_1, d_2, \dots$ may not be monotone.

(v) Provide lower bounds for $H(x, y, z; P_\lambda)$ of the expected order for other y, z not covered by Theorem 7.

Acknowledgements. The author thanks Bruce Berndt, Valery Nevzorov, Walter Philipp, Steven Portnoy, and Doug West for helpful conversations regarding probability estimates for uniform order statistics. The author also thanks Gérald Tenenbaum for several preprints of his work and for informing the author about the theorem of Rogers mentioned above, and thanks

Dimitris Koukoulopoulos for discussions which led to a simplification of the proof of Lemma 4.7. The author is grateful to his wife, Denka Kutzarova, for constant support and many helpful conversations about the paper. Much of this paper was written while the author enjoyed the hospitality of the Institute of Mathematics and Informatics, Bulgarian Academy of Sciences. Finally, the author acknowledges the referee for a thorough reading of the paper and for helpful suggestions.

This work was partially supported by National Science Foundation Grant DMS-0301083.

2. Preliminary lemmas

Further notation. $P^+(n)$ is the largest prime factor of n, $P^-(n)$ is the smallest prime factor of n. Adopt the conventions $P^+(1) = 0$ and $P^-(1) = \infty$. Also, $\omega(n)$ is the number of distinct prime divisors of n, $\Omega(n)$ is the number of prime power divisors of n, $\pi(x)$ is the number of primes $\le x$, $\tau(n)$ is the number of divisors of n. $\mathscr{P}(s,t)$ is the set of positive integers composed of prime factors p satisfying $s < p \le t$. Note that for all s,t we have $1 \in \mathscr{P}(s,t)$. $\mathscr{P}^*(s,t)$ is the set of square-free members of $\mathscr{P}(s,t)$.

We list a few estimates from prime number theory and sieve theory. The first is the Brun-Titchmarsh inequality and the second is a consequence of the Prime Number Theorem with classical de la Valée Poussin error term.

Lemma 2.1. *Uniformly in $x > y > 1$, we have $\pi(x) - \pi(x-y) \ll \frac{y}{\log y}$.*

Lemma 2.2. *For certain constants c_0, c_1,*

$$\sum_{p \le x} \frac{1}{p} = \log\log x + c_0 + O(e^{-c_1\sqrt{\log x}}) \qquad (x \ge 2).$$

The next result is a simple application of Brun's sieve (see [19]) together with the Prime Number Theorem with classical error term. Although not required for this paper, using bounds on the number of primes in short intervals, the best to date of which is [1], allows us to take Δ as small as $x^{0.525}$ in the lower bound in the next lemma.

Lemma 2.3. *Let $\Phi(x,z)$ be the number of integers $\le x$, all of whose prime factors are $> z$. If $1 < z^{1/100} \le \Delta \le x$, then*

$$\Phi(x,z) - \Phi(x-\Delta,z) \ll \frac{\Delta}{\log z}.$$

If $x \ge 2z$, z is sufficiently large and $xe^{-(c_1/2)\sqrt{\log x}} \le \Delta \le x$, then

$$\Phi(x,z) - \Phi(x-\Delta,z) \gg \frac{\Delta}{\log z}.$$

The second tool is crude but quite useful due to its uniformity. A proof may be found in Tenenbaum [44].

Lemma 2.4. *Let $\Psi(x,y)$ be the number of integers $\le x$, all of whose prime factors are $\le y$. Then, uniformly in $x \ge y \ge 2$,*

$$\Psi(x,y) \ll x \exp\{-\tfrac{\log x}{2\log y}\}.$$

Lemma 2.5. *Uniformly in $x>0$, $y \ge 2$ and $z \ge 1.5$, we have*

$$\sum_{\substack{n\ge x\\ n\in\mathscr{P}(z,y)}} \frac{1}{n} \ll \frac{\log y}{\log z} e^{-\frac{\log x}{4\log y}}, \tag{2.1}$$

$$\sum_{\substack{n\ge x\\ n\in\mathscr{P}(z,y)}} \frac{\log n}{n} \ll \frac{\log y \log(xy)}{\log z} e^{-\frac{\log x}{4\log y}}, \tag{2.2}$$

Proof. Without loss of generality we may assume that $x \ge 1$. Put $\alpha = \frac{1}{4\log y} \le \frac{2}{5}$. The result (2.1) is trivial unless $z<y$, in which case

$$\begin{aligned}\sum_{\substack{n\ge x\\ n\in\mathscr{P}(z,y)}} \frac{1}{n} &\le x^{-\alpha} \sum_{n\in\mathscr{P}(z,y)} n^{\alpha-1} = x^{-\alpha}\prod_{z<p\le y}\left(1+p^{\alpha-1}+p^{2(\alpha-1)}+\cdots\right)\\ &= x^{-\alpha}\exp\left\{\sum_{z<p\le y}\left(\frac{1}{p}+O\left(p^{-6/5}+\frac{\alpha\log p}{p}\right)\right)\right\}\\ &\ll x^{-\alpha}\frac{\log y}{\log z},\end{aligned}$$

where we used Lemma 2.2 in the final step. Now set

$$F(t) = \sum_{\substack{n>t\\ n\in\mathscr{P}(z,y)}} \frac{1}{n}.$$

By (2.1) and partial summation,

$$\begin{aligned}\sum_{\substack{n\ge x\\ n\in\mathscr{P}(z,y)}} \frac{\log n}{n} &= F(x)\log x + \int_x^\infty \frac{F(t)}{t}\,dt\\ &\ll \frac{\log y\,\log x}{\log z} e^{-\frac{\log x}{4\log y}} + \frac{\log y}{\log z}\int_x^\infty t^{-1-\frac{1}{4\log y}}\,dt\\ &\ll \frac{\log y\log(xy)}{\log z} e^{-\frac{\log x}{4\log y}}.\end{aligned}$$

□

We shall also need Stirling's formula

$$k! = \sqrt{2\pi k}(k/e)^k(1+O(1/k)), \tag{2.3}$$

although in most estimates weaker bounds will suffice.

Our last lemma is a consequence of Norton's bounds [31] for the partial sums of the exponential series. It is easily derived from Stirling's formula.

Lemma 2.6. *Suppose $0 \le h < m \le x$ and $m - h \ge \sqrt{x}$. Then*

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

3. Upper bounds outline

Initially, we bound $H(x,y,z)$ and $H_r(x,y,z)$ in terms of averages of the functions

$$L(a;\sigma) = \mathrm{meas}(\mathscr{L}(a;\sigma)), \quad \mathscr{L}(a;\sigma) = \{x : \tau(a, e^x, e^{x+\sigma}) \ge 1\}, \tag{3.1}$$

$$L_r(a;\sigma) = \mathrm{meas}(\mathscr{L}_r(a;\sigma)), \quad \mathscr{L}_r(a;\sigma) = \{x : \tau(a, e^x, e^{x+\sigma}) = r\}. \tag{3.2}$$

Here $\mathrm{meas}(\cdot)$ denotes Lebesgue measure. Both functions measure the global distribution of divisors of a. Before launching into the estimation of H and H_r, we list some basic inequalities for $L(a;\sigma)$.

Lemma 3.1. *We have*

(i) $L(a;\sigma) \le \min(\sigma\tau(a), \sigma + \log a)$;

(ii) *If* $(a,b) = 1$, *then* $L(ab;\sigma) \le \tau(b)L(a;\sigma)$;

(iii) *If* $(a,b) = 1$, *then* $L(ab;\sigma) \le L(a;\sigma + \log b)$;

(iv) *If* $\gamma \le \sigma$, *then* $L(a;\sigma) \le (\sigma/\gamma)L(a;\gamma)$;

(v) *If* $p_1 < \cdots < p_k$, *then*

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

Proof. Part (i) is immediate, since

$$\mathscr{L}(a;\sigma) = \bigcup_{d|a} [-\sigma + \log d, \log d) \subseteq [-\sigma, \log a).$$

Parts (ii) and (iii) follow from

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

Since $\mathscr{L}(a;\sigma)$ is a union of intervals of length σ, we obtain (iv). Combining parts (i) and (ii) with $a = p_1 \cdots p_j$ and $b = p_{j+1} \cdots p_k$ yields (v). □

Define

$$S(t;\sigma) = \sum_{P^+(a) \le te^\sigma} \frac{L(a;\sigma)}{a \log^2(t/a + P^+(a))}, \tag{3.3}$$

$$S_s(t;\sigma) = \sum_{P^+(a) \le te^\sigma} \frac{L_s(a;\sigma)}{a \log^2(t/a + P^+(a))}. \tag{3.4}$$

Lemma 3.2. *Suppose* $100 \le y \le \sqrt{x}$, $z = e^\eta y \le \min(x^{5/8}, y^{\log\log y})$ *and* $\eta \ge \frac{1}{\log y}$. *If* $x/\log^{10} z \le \Delta \le x$,

$$H(x,y,z) - H(x-\Delta,y,z) \ll \Delta \max_{y^{1/2} \le t \le x} S(t;\eta).$$

If in addition $y \ge y_0(r)$, *then*

$$H_r(x,y,z) \ll_r x \max_{\substack{1 \le s \le r \\ \nu(s) \le \nu(r)}} \max_{y^{1/2} \le t \le x} S_s(t;\eta).$$

Lemma 3.2 will be proved in Section 6. If $m < z/y$ then $\tau(n,y,z) \ge 1$ implies $\tau(nm,y,z^2/y) \ge 1$ and we expect (and prove) that $H(x,y,z)$ and $H(x,y,z^2/y)$ have the same order. Thus, for the problem of bounding $H(x,y,z)$, the prime factors of n below $z/y = e^\eta$ can essentially be ignored. For the problem of bounding $H_r(x,y,z)$, the prime factors of n less than z/y cannot be ignored and they play a different role in the estimation than the prime factors $> z/y$. In the next two lemmas, we estimate both $S(t;\sigma)$ and $S_s(t;\sigma)$ in terms of the quantity

$$S^*(t;\sigma) = \sum_{a \in \mathscr{P}^*(e^\sigma, te^\sigma)} \frac{L(a;\sigma)}{a \log^2(t^{3/4}/a + P^+(a))}. \tag{3.5}$$

Occasionally we will have need of the trivial lower bound

$$S^*(t;\sigma) \ge \frac{\sigma}{\log^2 t}, \tag{3.6}$$

obtained by taking the term $a = 1$ in (3.5).

Lemma 3.3. *Suppose* t *is large and* $0 < \sigma \le \log t$. *Then*

$$S(t;\sigma) \ll (1+\sigma) S^*(t;\sigma).$$

Particularly important in the estimation of $S_s(t;\sigma)$ is the distribution of the gaps between the first $r+1$ divisors of a, which ultimately depends on the power of 2 dividing r.

Lemma 3.4. *Suppose* $r \ge 1$, $C > 1$, $y \ge y_0(r,C)$, $z = e^\eta y$, $z \le x^{5/8}$ *and* $e^{100rC} y \le z \le y^C$. *Then*

$$H_r(x,y,z) \ll_{r,C} x (\log \eta)^{\nu(r)+1} \max_{y^{1/2} \le t \le x} S^*(t;\eta).$$

Lemmas 3.3 and 3.4 will be proved in Section 7.

To deal with the factor $\log^2(t^{3/4}/a + P^+(a))$ appearing in (3.5), define

$$T(\sigma, P, Q) = \sum_{\substack{a \in \mathscr{P}^*(e^\sigma, P) \\ a \ge Q}} \frac{L(a;\sigma)}{a}. \tag{3.7}$$

If $a \le t^{1/2}$ or $P^+(a) > t^{1/3000}$, then $\log^2(t^{3/4}/a + P^+(a)) \gg \log^2 t$. Otherwise, $e^{e^{g-1}} < P^+(a) \le e^{e^g}$ for some integer g satisfying $e^\sigma \le e^{e^g} \le t^{1/1000}$. Thus we have

$$S^*(t;\sigma) \ll \frac{T(\sigma, te^\sigma, 1)}{\log^2 t} + \sum_{\substack{g \in \mathbb{Z}, g \ge 1 \\ e^\sigma \le e^{e^g} \le t^{1/1000}}} e^{-2g} T(\sigma, e^{e^g}, t^{1/2}). \tag{3.8}$$

We break up the sum in $T(\sigma, P, Q)$ according to the value of $\omega(a)$ and define

$$T_k(\sigma, P, Q) = \sum_{\substack{a \in \mathscr{P}^*(e^\sigma, P) \\ a \ge Q \\ \omega(a) = k}} \frac{L(a;\sigma)}{a}.$$

Note that the definition of k given here is slightly different from that mentioned in the heuristic argument of subsection 1.5, but usually differs only by $O(1)$. By Lemma 2.2 and part (v) of Lemma 3.1, $T_k(\sigma, P, Q)$ will be bounded in terms of

$$U_k(v;\alpha) = \int_{R_k} \min_{0 \le j \le k} 2^{-j} \left(2^{v\xi_1} + \cdots + 2^{v\xi_j} + \alpha\right) d\boldsymbol{\xi}, \tag{3.9}$$

where

$$R_k = \{\boldsymbol{\xi} \in \mathbb{R}^k : 0 \le \xi_1 \le \cdots \le \xi_k \le 1\}.$$

For convenience, let $U_0(v;\alpha) = \alpha$.

Lemma 3.5. *Suppose $P \ge 100$, $0 < \sigma < \log P$, and $Q \ge 1$. Let*

$$v = \left\lceil \frac{\log\log P - \max(0, \log\sigma)}{\log 2} \right\rceil$$

and suppose $0 \le k \le 10v$. Then

$$T_k(\sigma, P, Q) \ll e^{-\frac{\log Q}{\log P}} (\sigma + 1)(2v\log 2)^k U_k(v; \min(1, \sigma)).$$

Lemma 3.5 will be proved in Section 8. As a rough heuristic, $2^{v\xi_1} + \cdots + 2^{v\xi_j} \ll 2^{v\xi_j}$ most of the time. Thus, bounding $U_k(v;\alpha)$ boils down to determining the distribution in R_k of the function

$$F(\boldsymbol{\xi}) = \min_{1 \le j \le k} (\xi_j - j/v).$$

The numbers $\xi_1, \ldots, \xi_k$ can be regarded as independent uniformly distributed random variables on $[0, 1]$, relabeled to have the above ordering, and are known as *uniform order statistics*. Making this heuristic precise, and using results about the distribution of uniform order statistics from Section 11, leads to the next result, which will be proved in Section 13.

LEMMA 3.6. *Suppose k, v are integers with $0 \le k \le 10v$ and $0 < \alpha \le 1$. Then*

$$U_k(v;\alpha) \ll \frac{\alpha \min\big(k+1, (1+|v-k-\frac{\log\alpha}{\log 2}|^2)\log(2/\alpha)\big)}{(k+1)!(\alpha 2^{k-v}+1)}.$$

Notice that, as a function of k, the bound in Lemma 3.6 undergoes a change of behavior at $k = \left\lfloor v - \frac{\log\alpha}{\log 2}\right\rfloor$. It is now straightforward to give a relatively simple upper bound for $T(\sigma, P, Q)$.

LEMMA 3.7. *Suppose P is sufficiently large, $Q \ge 1$, and*

$$(\log P)^{-1} \le \sigma < \log P.$$

Define $\theta = \theta(\sigma, P)$ and $\nu = \nu(\sigma, P)$ by $\sigma = (\log P)^{-\theta}$ and $\theta = \log 4 - 1 - \nu(\log\log P)^{-1/2}$ (these quantities are related to those in (1.2)*). Then*

$$T(\sigma,P,Q) \ll \begin{cases} e^{-\frac{\log Q}{\log P}} \dfrac{\sigma^{\delta-1}(\log P)^{2-\delta}}{(\log\frac{\log P}{1+\sigma}+1)^{3/2}} & \sigma \ge 1 \\[2ex] e^{-\frac{\log Q}{\log P}} \dfrac{(\log P)^{2-G(\theta)}\log(2/\sigma)}{\max(1,\nu)\log\log P} & \sigma < 1. \end{cases}$$

Proof. Define v as in the statement of Lemma 3.5 and set $\alpha = \min(1, \sigma)$. Put $\gamma = e^{-\frac{\log Q}{\log P}}$. By Lemmas 3.5 and 3.6, when $0 \le k \le 10v$,

$$\begin{aligned} T_k(\sigma,P,Q) &\ll \gamma\alpha(\sigma+1)Z_k \ll \gamma\sigma Z_k, \\ Z_k &= \frac{\min(k+1,(1+|v-k-\frac{\log\alpha}{\log 2}|^2)\log(2/\alpha))}{(k+1)!(\alpha 2^{k-v}+1)}(2v\log 2)^k. \end{aligned} \tag{3.10}$$

Put $k_1 = \left\lfloor v - \frac{\log\alpha}{\log 2}\right\rfloor$ and note that $v \le k_1 \le 2v$. Now,

$$\sum_{k_1\le k\le 10v} Z_k \ll \log(2/\alpha)\sum_{b\ge 0}\frac{b^2+1}{2^b}\frac{(2v\log 2)^{k_1+b}}{(k_1+b+1)!} \ll \frac{\log(2/\alpha)(2v\log 2)^{k_1}}{(k_1+1)!}. \tag{3.11}$$

By $L(a;\sigma) \le 2^{\omega(a)}\sigma$,

$$\begin{aligned} \sum_{k\ge 10v} T_k(\sigma,P,Q) &\le Q^{-1/\log P}\sigma \sum_{k\ge 10v} 2^k \sum_{\substack{a\in\mathscr{P}^*(e^\sigma,P)\\ \omega(a)=k}} \frac{1}{a^{1-1/\log P}} \\ &\le \gamma\sigma\sum_{k\ge 10v}\frac{2^k}{k!}\Bigg(\sum_{e^\sigma<p\le P}\frac{1}{p^{1-1/\log P}}\Bigg)^k \end{aligned}$$

$$= \gamma\sigma \sum_{k \ge 10v} \frac{(2v\log 2 + O(1))^k}{k!}$$
$$\ll \gamma\sigma \frac{(2v\log 2)^{10v}}{(10v)!}$$
$$\ll \gamma\sigma \frac{(2v\log 2)^{k_1}}{(k_1+1)!}.$$

Together with (3.10) and (3.11), we conclude that

$$\sum_{k \ge k_1} T_k(\sigma, P, Q) \ll \frac{\gamma\sigma \log(2/\alpha)(2v\log 2)^{k_1}}{(k_1+1)!}. \tag{3.12}$$

Suppose that $\theta \le \frac{1}{3}$, so that $k_1 \le \frac{4}{3}v$. Since $\frac{4}{3} < 2\log 2$, (3.10) implies

$$\sum_{0 \le k \le k_1} Z_k \ll \log(2/\alpha) \sum_{0 \le k \le k_1} (k_1 - k + 1)^2 \frac{(2v\log 2)^k}{(k+1)!} \ll \frac{\log(2/\alpha)(2v\log 2)^{k_1}}{(k_1+1)!}.$$

Combined with (3.10), (3.12) and Stirling's formula, this gives

$$\begin{aligned} T(\sigma, P, Q) &\ll \frac{\gamma\sigma \log(2/\alpha)(2v\log 2)^{k_1}}{(k_1+1)!} \\ &\ll \frac{\gamma\sigma\log(2/\alpha)}{v^{3/2}} \left(\frac{2ev\log 2}{k_1}\right)^{k_1} \qquad (\theta \le \tfrac{1}{3}). \end{aligned} \tag{3.13}$$

When $\sigma \ge 1$, we have $\theta \le 0$, $\alpha = 1$, $k_1 = v$, and

$$(2e\log 2)^v \asymp \left(\frac{\log P}{\sigma}\right)^{2-\delta},$$

and so the lemma follows in this case. We also have

$$v = \frac{\log\log P}{\log 2} + O(1), \qquad k_1 = (1+\theta)\frac{\log\log P}{\log 2} + O(1) \qquad (0 \le \theta \le 1)$$

and

$$\left(\frac{2ev\log 2}{k_1}\right)^{k_1} \asymp (\log P)^{2+\theta-G(\theta)} \qquad (0 \le \theta \le \log 4 - 1).$$

Thus, if $0 \le \theta \le \frac{1}{3}$, then $\nu \asymp (\log\log P)^{1/2}$ and the lemma follows from (3.13). If $\frac{1}{3} \le \theta \le 1$, (3.10) and Lemma 2.6 give

$$\sum_{0 \le k \le k_1} Z_k \ll \sum_{0 \le k \le k_1} \frac{(2v\log 2)^k}{k!} \ll \begin{cases} e^{2v\log 2}, & k_1 \ge 2v\log 2 - \sqrt{v} \\ \dfrac{\sqrt{v}(2v\log 2)^{k_1}}{\nu k_1!}, & k_1 < 2v\log 2 - \sqrt{v} \end{cases}$$
$$\ll \frac{(\log P)^{2+\theta-G(\theta)}\log(2/\sigma)}{\max(1,\nu)\log\log P}.$$

Together with (3.10) and (3.12), this proves the lemma in the final case. □

Lemma 3.8. *Suppose t is large and $\sigma \ge (\log t)^{-1/2}$. Put $\theta = \theta(\sigma, t)$ and $\nu = \nu(\sigma, t)$. Then*

$$S^*(t,\sigma) \ll \begin{cases} \dfrac{\sigma^{\delta-1}(\sigma+\log t)^{2-\delta}}{(\log t)^2(\log(\sigma+\log t)-\log\sigma+1)^{3/2}}, & \sigma \ge 1 \\[2ex] \dfrac{\log(2/\sigma)}{(\log\log t)\max(1,\nu)(\log t)^{G(\theta)}}, & \sigma \le 1. \end{cases}$$

Proof. First suppose $\sigma \ge 1$. By (3.8) and Lemma 3.7, writing $g = \lfloor \log\log t \rfloor - \ell$ gives

$$\begin{aligned} S^*(t;\sigma) \ll \frac{\sigma^{\delta-1}}{(\log t)^{\delta}} \Bigg[\left(\frac{\sigma+\log t}{\log t} \right)^{2-\delta} & \frac{1}{(\log(\sigma+\log t)-\log\sigma+1)^{3/2}} \\ + \sum_{1\le \ell \le \log\log t - \log\sigma} & \frac{e^{\delta\ell}}{e^{e^{\ell-1}}(\log\log t - \log\sigma + 1 - \ell)^{3/2}} \Bigg]. \end{aligned}$$

The sum on ℓ is empty if $\sigma > \log t$. Otherwise, the sum on ℓ is dominated by terms with $\ell \ll 1$, and this proves the lemma in this case.

Suppose that $\sigma < 1$. By Lemma 3.7, the first term in (3.8) is

$$\ll \frac{\log(2/\sigma)}{\max(1,\nu)(\log t)^{G(\theta)}\log\log t}.$$

We use Lemma 3.7 when $e^{-g} \le \sigma$. The contribution of these terms (if any) in (3.8) is

$$\begin{aligned} &\ll \log(2/\sigma) \sum_{\log\frac{1}{\sigma} < g \le \log\log t} \frac{e^{-gG(-(\log\sigma)/g)}}{e^{(e^{-g-1}\log t)} g \max(1, \sqrt{g}(1-\log 4 - \frac{\log\sigma}{g}))} \\ &\ll \frac{\log(2/\sigma)}{\max(1,\nu)(\log t)^{G(\theta)}\log\log t}. \end{aligned}$$

When $g < \log(1/\sigma) \le \frac{1}{2}\log\log t$, Lemma 3.5 gives

$$T_k(\sigma, e^{e^g}, t^{1/2}) \ll e^{-\frac{1}{2}\sqrt{\log t}}(2v\log 2)^k U_k(v;\sigma) \le e^{-\frac{1}{2}\sqrt{\log t}}(2v\log 2)^k \sigma/k!,$$

where $v = \frac{g}{\log 2} + O(1)$. Summing on k and g yields

$$\sum_{g<\log\frac{1}{\sigma}} e^{-2g} T(\sigma, e^{e^g}, t^{1/2}) \ll \sum_{g \le \frac{1}{2}\log\log t} \sigma e^{-\frac{1}{2}\sqrt{\log t}} \ll \exp\{-\frac{1}{3}(\log t)^{1/2}\},$$

which is negligible compared to the contribution of the terms in (3.8) with $g \ge \log(1/\sigma)$. $\square$

For fixed σ, $\theta(\sigma,t)$ is decreasing, $\nu(\sigma,t)$ is increasing and $G(\theta)/\theta$ is increasing, as functions of t. Thus, we have the following.

LEMMA 3.9. *Suppose y is large and $\eta \ge (\log y)^{-0.4}$. Then*

$$\max_{t \ge y^{1/2}} S^*(t;\eta) \ll \begin{cases} \dfrac{\eta^{\delta-1}(\eta+\log y)^{2-\delta}}{(\log y)^2(\log(\eta+\log y)-\log\eta+1)^{3/2}}, & \eta \ge 1 \\ \\ \dfrac{\log(2/\eta)}{(\log\log y)\max(1,\nu(\eta,y))(\log y)^{G(\theta(\eta,y))}}, & \eta \le 1. \end{cases}$$

4. Lower bounds outline

As with the upper bounds, we initially bound $H(x,y,z)$ in terms of sums over $L(a;\sigma)$ and bound $H_r(x,y,z)$ in terms of sums over $L_s(a;\sigma)$ (but only for $s=r$). The initial bounds are similar to those in Lemma 3.2.

LEMMA 4.1. *Suppose $y_0 \le y < z = e^\eta y$, $\frac{1}{\log^{20} y} \le \eta \le \frac{\log y}{100}$, $y \le \sqrt{x}$ and $x/\log^{10} z \le \Delta \le x$. Then*

$$\begin{aligned} H(x,y,z) - H(x-\Delta,y,z) &\ge H^*(x,y,z) - H^*(x-\Delta,y,z) \\ &\gg \frac{\Delta}{\log^2 y} \sum_{\substack{a \le y^{1/8} \\ \mu^2(a)=1}} \frac{L(a;\eta)}{a}. \end{aligned}$$

Suppose $r \ge 1$, $0 < c \le \frac18$, $C > 0$, $y_0(r,c,C) \le y < z = e^\eta y$, $\frac{1}{\log^2 y} \le \eta \le C\log y$ and $z \le x^{1/2-c}$. Then

$$H_r(x,y,z) \gg_{r,c,C} \frac{x}{\log^2 y} \sum_{a \le y^{2c}} \frac{L_r(a;\eta)}{a}.$$

Lemma 4.1 will be proved in Section 6. Both $L(a;\sigma)$ and $L_r(a;\sigma)$ may be bounded below in terms of the function

$$I(n;\sigma) = |\{d|n : \tau(n, de^{-\sigma}, de^{\sigma}) = 1\}|. \tag{4.1}$$

Introduced by Tenenbaum [43], $I(n;\sigma)$ counts σ-*isolated* divisors of n.

In the first part of Lemma 4.1, take square-free $a = h'h$, where $h' \le z/y \le y^{1/100}$ and $P^-(h) > z/y$. Clearly

$$L(h'h;\eta) \ge L(h;\eta) \ge \eta I(h;\eta),$$

and summing over g we obtain the following.

LEMMA 4.2. *Suppose $y_0 \le y < z = e^\eta y$, $\frac{1}{\log^{20} y} \le \eta \le \frac{\log y}{100}$, $y \le \sqrt{x}$ and $\frac{x}{\log^{10} z} \le \Delta \le x$. Then*

$$H^*(x,y,z) - H^*(x-\Delta,y,z) \gg \frac{\eta(1+\eta)\Delta}{\log^2 y} \sum_{\substack{h \le y^{1/10} \\ P^-(h) > z/y \\ \mu^2(h)=1}} \frac{I(h;\eta)}{h}.$$

We follow two methods for bounding $H_r(x,y,z)$ from below, the first useful for $z \ll y$ and the second useful for large z.

LEMMA 4.3. *Suppose $r \ge 1$, $0 < c' \le \frac{1}{8}$, $y_0(r,c') \le y < z = e^\eta y \le x^{1/2-c'}$ and $\frac{1}{\log^2 y} \le \eta \le \frac{c' \log y}{10r}$. Then*

$$H_r(x,y,z) \gg_{r,c'} \frac{\eta^r x}{(\log y)^{r+1}} \sum_{a \le y^{c'/100r}} \frac{I(a;\eta)^r}{a}.$$

Lemma 4.3 and its proof are essentially taken from Lemme 4 of Tenenbaum [43]. The main difference is the upper limit of allowable z: Lemme 4 of [43] requires $z \le x^{\frac{1}{r+1}-c}$.

In the second method, the prime factors of a which are $< z/y$ play a special role as in Lemma 3.4.

LEMMA 4.4. (i) *Suppose $r \ge 1$, $C > 0$, $0 < c' \le \frac{1}{8}$, $y_0(r,c',C) \le y < z = e^\eta y \le x^{1/2-c'}$ and $1000r \cdot 3^{2r} \le \eta \le C \log y$. Then*

$$H_r(x,y,z) \gg_{r,c',C} \frac{\eta(\log \eta)^{\nu(r)+1} x}{\log^2 y} \sum_{\substack{h \le y^{c'} \\ P^-(h) > e^{2\eta}}} \frac{I(h;2\eta)}{h}.$$

(ii) *If $r \ge 1$, $0 < c \le \frac{1}{8}$, $y \ge y_0(r,c)$ and $y^2 \le z \le x^{1-c}/y$, then*

$$H_r(x,y,z) \gg_{r,c} \frac{x(\log\log y)^{\nu(r)+1}}{\log z}.$$

Lemmas 4.3 and 4.4 will be proved in Section 9. The number of isolated divisors of a number can be easily bounded from below in terms of

$$W(a;\sigma) = |\{(d_1,d_2) : d_1|a, d_2|a, |\log(d_1/d_2)| \le \sigma\}|. \tag{4.2}$$

This function, introduced by Hall [21], is essential in the study of the propinquity of divisors (see also [23], [28], [29], Chapters 4 and 5 of [24], [33], and [45]). The following lemma is similar to Lemme 5 of Tenenbaum [43].

LEMMA 4.5. *There exists $I(a;\sigma)$ such that*

$$I(a;\sigma)^r \ge 2^{-r}\tau(a)^{r-1}(3\tau(a) - 2W(a;\sigma)).$$

Proof. For each divisor d of a not counted by $I(a;\sigma)$ there is at least one other divisor d' satisfying $d/e^\sigma \le d' \le de^\sigma$, so that the pair (d,d') is counted by $W(a;\sigma)$. Thus

$$W(a;\sigma) \ge \tau(a) + (\tau(a) - I(a;\sigma)) = 2\tau(a) - I(a;\sigma).$$

The lemma is trivial when $W(a;\sigma) \ge \frac{3}{2}\tau(a)$. Otherwise,

$$I(a;\sigma)^r \ge (2\tau(a) - W(a;\sigma))^r \ge \left(\frac{\tau(a)}{2}\right)^{r-1} (\tfrac{3}{2}\tau(a) - W(a;\sigma)). \qquad \square$$

With Lemma 4.5, lower bounds for $H(x,y,z)$ and $H_r(x,y,z)$ are obtained via upper bounds on sums over $W(a;\sigma)/a$. Such upper bounds are achieved by partitioning the primes into sets $D_1, D_2, \ldots$ and separately considering numbers a with a fixed number of prime factors in each interval D_j.

Each set D_j will consist 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 inductively define λ_j for $j \ge 1$ as the largest prime so that

$$\sum_{\lambda_{j-1} < p \le \lambda_j} \frac{1}{p} \le \log 2. \tag{4.3}$$

For example, $\lambda_1 = 2$ and $\lambda_2 = 7$. Write $\lambda_j = \exp\{2^{\mu_j}\}$.

Lemma 4.6. *There are constants c_3, c_4 so that $|\mu_j - j - c_3| \le c_4 2^{-j}$ for all $j \ge 0$.*

Proof. Clearly $\lambda_j \to \infty$ as $j \to \infty$. By (4.3) and Lemma 2.2 with crude error term,

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

Thus, for large j, $\log \lambda_j \ge 1.9 \log \lambda_{j-1}$ and hence $\sum_j 1/\log \lambda_{j-1}$ converges. Now, $\mu_j = j + O(1)$ and for $r > s \ge 1$

$$\mu_r - \mu_s = r - s + O\left(\sum_{j \ge s} 2^{-\mu_{j-1}} \right) = r - s + O(2^{-s}). \tag{4.4}$$

Therefore, the sequence $(\mu_j - j)$ is a Cauchy sequence converging to some value c_3, and $|\mu_j - j - c_3| = O(2^{-j})$. $\square$

For a vector $\mathbf{b} = (b_1, \ldots, b_h)$ of non-negative integers, let $\mathscr{A}(\mathbf{b})$ be the set of square-free integers a composed of exactly b_j prime factors in D_j for each j. Denote $k = b_1 + \cdots + b_h$. For the remainder of this section, M will be a sufficiently large absolute constant, which we take to be an even integer.

Lemma 4.7. *Suppose $\sigma > 0$, $\mathbf{b} = (b_1, \ldots, b_h)$ and define $m = \min\{j : b_j \ge 1\}$. If $\sigma < 1$, further assume that $m \ge M$ and $b_j \le 2^{j/2}$ for each j. Then*

$$\sum_{a \in \mathscr{A}(\mathbf{b})} \frac{W(a;\sigma)}{a} \le \frac{(2\log 2)^k}{b_m! \cdots b_h!} \left[1.01 + 2^{c_5} \sigma \sum_{j=m}^{h} 2^{-j+b_m+\cdots+b_j} \right],$$

where c_5 is an absolute constant.

We next apply Lemma 4.7 for many vectors $\mathbf{b}$.

Lemma 4.8. *Suppose* $0 < \alpha \le 1$, $y \ge y_0(\alpha)$ *and* $0 < \sigma \le 2^{-2M-1/\alpha} \log y$. *Define*

$$v = \left\lfloor \frac{\log\log y - \max(0, \log\sigma)}{\log 2} - 2M - 1/\alpha + 1 \right\rfloor,$$
$$s = M + \max\left(0, \left\lfloor \frac{\log\sigma}{\log 2} \right\rfloor\right) - c_5 - 10 - \frac{\log\sigma}{\log 2}.$$

Suppose $k \ge M+1$. *Then, for some subset* $\mathscr{A}$ *of the squarefree integers* $a \le y^\alpha$ *satisfying* $P^-(a) > e^\sigma$ *and* $\omega(a) = k$, *we have*

$$\sum_{a\in\mathscr{A}} \frac{3\tau(a) - 2W(a;\sigma)}{a} \ge \frac{1}{3}(2v\log 2)^k \operatorname{Vol}(Y_k(s,v)),$$

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

(i) $0 \le \xi_1 \le \cdots \le \xi_k < 1$;

(ii) *For* $1 \le i \le \sqrt{k-M}$, $\xi_{M+i^2} > i/v$ *and* $\xi_{k+1-(M+i^2)} < 1 - i/v$;

(iii) $\sum_{j=1}^k 2^{j-v\xi_j} \le 2^s$.

Condition (ii) in the definition of $Y_k(s,v)$ is very mild and does not affect the volume very much. It arises from taking numbers $a \in \mathscr{A}$ which have neither too many small prime factors nor too many large prime factors. Lemmas 4.7 and 4.8 will be proved in Section 10. The volume of $Y_k(s,v)$ can be estimated using bounds on uniform order statistics (§11). Once the volume has been bounded from below, we are in position to complete the lower bounds in Theorems 1–5.

Lemma 4.9. *Suppose* $v \ge 1$, $10M \le k \le 100(v-1)$, $s \ge M/2 + 1$ *and* $0 \le k - v \le s - M/3 - 1$. *Then*

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

Lemma 4.9 will be proved in Section 12.

5. Proof of Theorems 1, 2, 3, 4 and 5

Suppose throughout that $z \ge y+1$ and $y \ge y_0$ (Theorem 1 is trivial if $y < y_0$).

Upper bounds in Theorems 1, 2 *and* 3 *when* $y \le \sqrt{x}$. If $0 < \eta \le 1$ and $\Delta \ge \sqrt{x}$, then

$$\begin{aligned} H(x,y,z) - H(x-\Delta,y,z) &\le \sum_{y<d\le z} \left(\frac{\Delta}{d} + 1 \right) \\ &\ll \eta\Delta + (z-y) \\ &\ll \eta\Delta. \end{aligned}$$

This proves the upper bounds in the three theorems when $z \le z_0(y)$. For $z_0(y) \le z \le y^{1.001}$, the desired bounds follow from Lemmas 3.2, 3.3 and 3.9. When $\beta \gg 1$, our upper bound coincides with that of Theorem T1 (ii). When $z \ge y^{1.001}$, the trivial bound $H(x,y,z) - H(x-\Delta,y,z) \le \Delta + 1$ suffices.

Lower bounds in Theorems 1, 2 *and* 3 *when* $y \le \sqrt{x}$. Assume $\frac{x}{\log^{10} z} \le \Delta \le x$ for the estimation of $H(x,y,z) - H(x-\Delta,y,z)$ and $\frac{x}{\log y} \le \Delta \le x$ in the estimation of $H^*(x,y,z) - H^*(x-\Delta,y,z)$.

If $0 < \eta \le \frac{1}{\log^{20} y}$, then

$$\begin{aligned} H(x,y,z) - H(x-\Delta,y,z) &\ge \sum_{y<d\le z} \left\lfloor \frac{x}{d} \right\rfloor - \left\lfloor \frac{x-\Delta}{d} \right\rfloor \\ &\quad - \sum_{y<d_1<d_2\le z} \left\lfloor \frac{x}{\mathrm{lcm}[d_1,d_2]} \right\rfloor - \left\lfloor \frac{x-\Delta}{\mathrm{lcm}[d_1,d_2]} \right\rfloor \\ &\ge \Delta \left(\sum_{y<d\le z} \frac{1}{d} - \sum_{y<d_1<d_2\le z} \frac{1}{\mathrm{lcm}[d_1,d_2]} \right) - 2(z-y+1)^2. \end{aligned}$$

Let $m = (d_1,d_2)$, so that $m \le z-y$. Write $d_1 = t_1 m$, $d_2 = t_2 m$. Then

$$\begin{aligned} H(x,y,z) - H(x-\Delta,y,z) &\ge \Delta \left(\sum_{y<d\le z} \frac{1}{d} - \sum_{m\le z-y} \frac{1}{m} \sum_{\frac{y}{m}<t_1<t_2\le \frac{z}{m}} \frac{1}{t_1 t_2} \right) \\ &\quad - O(\eta^2 y^2) \\ &= \Delta \sum_{y<d\le z} \frac{1}{d} - O(\Delta\eta^2 \log y + \eta^2 x) \\ &= \Delta \sum_{y<d\le z} \frac{1}{d} - O(\eta\Delta(\log y)^{-10}). \end{aligned}$$

The sum on d is $\gg \eta$, and we conclude the lower bound in Theorem 1 (v) and Theorem 2 for this range of y, z.

Now suppose $g > 0$, $y \ge y_0(g)$, $0 < \eta \le \frac{1}{\log^{20} y}$ and there are $\ge g(z-y)$ square-free integers in $(y,z]$. By a theorem of Filaseta and Trifonov [13], this last condition holds unconditionally with $g = \frac{1}{2}$ provided $z \ge y + Ky^{1/5}\log y$

for a large constant K. We obtain

$$\begin{aligned} H^*(x,y,z) - H^*(x-\Delta,y,z) &\ge \sum_{y<d\le z} \sum_{\substack{x-\Delta<e\le x\\ d|e}} \mu^2(e) - \sum_{y<d_1<d_2\le z} \frac{x}{\operatorname{lcm}[d_1,d_2]} \\ &\ge \sum_{y<d\le z} \mu^2(d) \sum_{\substack{\frac{x-\Delta}{d}<f\le \frac{x}{d}\\ (f,d)=1}} \mu^2(f) - O(\eta^2 x\log y). \end{aligned}$$

A simple elementary argument yields

$$\sum_{\substack{f\le w\\ (d,f)=1}} \mu^2(f) = C_d w + O\left(w^{1/2}\tau(d)\right),$$

where

$$C_d = \frac{\phi(d)}{d}\prod_{p\nmid d}(1-1/p^2).$$

Thus,

$$H^*(x,y,z) - H^*(x-\Delta,y,z) \gg \frac{\Delta}{y^2}\sum_{y<d\le z}\mu^2(d)\phi(d) - O(\eta\Delta(\log y)^{-18}).$$

Now apply the estimate

$$\sum_{n\le x}\frac{1}{\phi(n)} = C_1\log x + C_2 + O\left(\frac{(\log x)^{2/3}}{x}\right) \tag{5.1}$$

due to Sitaramachandra Rao [35], where C_1, C_2 are certain constants (Landau had in 1900 proved a weaker version with error term $O(\frac{\log x}{x})$). By the Cauchy-Schwarz inequality and our assumption,

$$\sum_{y<d\le z}\mu^2(d)\phi(d) \ge \left(\sum_{y<d\le z}\mu^2(d)\right)^2\left(\sum_{y<d\le z}\frac{1}{\phi(d)}\right)^{-1} \gg \eta y^2.$$

We conclude that

$$H^*(x,y,z) - H^*(x-\Delta,y,z) \gg \eta\Delta,$$

which completes the proof of the lower bound in Theorem 3 in this case.

Next, suppose $\frac{1}{\log^{20} y} \le \eta \le \frac{1}{100}$ and define β, ξ by (1.2). Let $\sigma = \eta$, $0<\alpha\le 1$, $g\ge 1$ and $y \ge y_0(\alpha, g)$. In Lemma 4.8, we have

$$\begin{aligned} v &= \left\lfloor \frac{\log\log y}{\log 2} - 2M - 1/\alpha + 1\right\rfloor, \\ s &= M + c_3 - 21 - \frac{\log\eta}{\log 2} \ge \frac{M}{2} + 1 - \frac{\log\eta}{\log 2}. \end{aligned}$$

We will apply Lemmas 4.8 and 4.9 with all k satisfying

$$(1+\beta/100)v \le k \le \min(1+\beta, \log 4)v.$$

This includes at least one value of k since $\frac{\log 100}{\log\log y} \le \beta \le 20$. Also, by (1.2),

$$k - v \le \beta v = \frac{-\log \eta}{\log\log y} v \le s - M/3 - 1,$$

so that all hypotheses of Lemma 4.9 are satisfied. For each such k we obtain

$$\sum_{a\in\mathscr{A}} \frac{3\tau(a) - 2W(a;\eta)}{a} \gg_{g,\alpha} \beta \frac{(2v\log 2)^k}{k!},$$

for some subset $\mathscr{A}$ of the squarefree integers $a \le y^\alpha$ with $\omega(a) = k$ and $P^-(a) > e^\eta$. By Lemma 4.5,

$$\sum_{\substack{a\le y^\alpha \\ \omega(a)=k \\ \mu^2(a)=1}} \frac{I(a;\eta)^g}{a} \gg_{g,\alpha} \beta \frac{(2^g v\log 2)^k}{k!}. \tag{5.2}$$

When $g = 1$, Lemma 2.6 gives

$$\frac{\eta}{\log^2 y} \sum_{\substack{a\le y^\alpha \\ \mu^2(a)=1}} \frac{I(a;\eta)}{a} \gg_\alpha \frac{\beta}{\max(1,-\xi)(\log y)^{G(\beta)}}. \tag{5.3}$$

By Lemma 4.2 and (5.3) with $\alpha = \frac{1}{10}$, we obtain

$$\begin{aligned} H(x,y,z) - H(x-\Delta,y,z) &\ge H^*(x,y,z) - H^*(x-\Delta,y,z) \\ &\gg \frac{\beta\Delta}{\max(1,-\xi)(\log y)^{G(\beta)}}. \end{aligned} \tag{5.4}$$

Next, let $0 < \alpha \le 1$, $\gamma = 2^{-20M-1/\alpha}$, and suppose that $\frac{1}{100} \le \eta \le \gamma \log y$. Let $g \ge 1$, $y \ge y_0(\alpha, g)$ and assume $\sigma = \eta$ or $\sigma = 2\eta$. In Lemmas 4.8 and 4.9, $v \ge 18M$ and $s \ge M/2 + 1$ if M is large enough. Using the single term $k = v$, we have (note $e^\eta = y^u$)

$$\sum_{a\in\mathscr{A}} \frac{3\tau(a) - 2W(a;\sigma)}{a} \gg_{g,\alpha} \frac{(2v\log 2)^v}{v\cdot v!} \gg \frac{u^{\delta-2}}{(-\log u)^{3/2}},$$

for a subset $\mathscr{A}$ of the squarefree integers $a \le y^\alpha$ satisfying $\omega(a) = k$ and $P^-(a) > e^\sigma$. Lemma 4.5 then gives

$$\sum_{\substack{a\le y^\alpha \\ P^-(a)>e^\sigma \\ \mu^2(a)=1}} \frac{I(a;\sigma)^g}{a} \gg_{g,\alpha} \frac{2^{v(g-1)}u^{\delta-2}}{(-\log u)^{3/2}} \gg_g \frac{u^{\delta-1-g}}{(-\log u)^{3/2}}. \tag{5.5}$$

By Lemma 4.2 and (5.5) with $g=1$ and $\alpha=\frac{1}{10}$,

$$H^*(x,y,z)-H^*(x-\Delta,y,z)\gg\Delta\frac{u^\delta}{(-\log u)^{3/2}}. \tag{5.6}$$

Finally, if $z\ge y^{1+\gamma}$, then by (5.6)

$$H(x,y,z)-H(x-\Delta,y,z)\ge H(x,y,y^{1+\gamma})-H(x-\Delta,y,y^{1+\gamma})\gg\Delta$$

and

$$H^*(x,y,z)-H^*(x-\Delta,y,z)\ge H^*(x,y,y^{1+\gamma})-H^*(x-\Delta,y,y^{1+\gamma})\gg\Delta.$$

Corollary 1 also follows in the case $y_i\le\sqrt{x_i}$ $(i=1,2)$.

Proof of Theorem 1 (vi). Assume throughout that $y>\sqrt{x}$ and $y+1\le z\le x$. If $\frac{x}{y}<\frac{x}{z}+1$ and $y<d_1<d_2\le z$, then

$$\operatorname{lcm}[d_1,d_2]=\frac{d_1d_2}{(d_1,d_2)}>\frac{y(y+(d_1,d_2))}{(d_1,d_2)}\ge y+\frac{y^2}{z-y}\ge x.$$

Hence

$$H(x,y,z)=\sum_{y<d\le z}\left\lfloor\frac{x}{d}\right\rfloor\asymp\eta x.$$

Now assume $\frac{x}{y}\ge\frac{x}{z}+1$. If $y>x/y_0$, then $H(x,y,z)\gg\lfloor z\rfloor-\lfloor y\rfloor\gg x$ since $z\ge\frac{xy}{x-y}\ge\frac{y_0}{y_0-1}y$. Thus

$$H(x,y,z)\asymp x\asymp H(x,\tfrac{x}{z},\tfrac{x}{y}).$$

Next suppose $\sqrt{x}<y\le x/y_0$ and $\eta\le\frac{1}{\log^2(x/y)}$. Here,

$$H(x,y,z)=\sum_{y<d\le z}\left\lfloor\frac{x}{d}\right\rfloor+O\left(\sum_{\substack{y<d_1<d_2\le z\\ \operatorname{lcm}[d_1,d_2]\le x}}\frac{x}{\operatorname{lcm}[d_1,d_2]}\right).$$

The sum on d is $\asymp\eta x$. Writing $m=(d_1,d_2)$, the big-O term is

$$\ll\sum_{y^2/x<m\le z-y}\frac{1}{m}\sum_{\frac{y}{m}<t_1<t_2\le\frac{z}{m}}\frac{1}{t_1t_2}\ll\eta^2\log(x/y)\ll\frac{\eta}{\log(x/y)}.$$

Since $\log(x/y)\asymp\log(x/z)$,

$$H(x,y,z)\asymp\eta x\asymp H(x,\tfrac{x}{z},\tfrac{x}{y})$$

in this case.

Lastly, suppose $\eta\ge\frac{1}{\log^2(x/y)}$. Partition $(\frac{x}{\log^2(x/y)},x]$ into intervals $(x_1,x_2]$, where

$$x_2-x_1\in\left[\frac{x_2}{\log^3(x/y)},\frac{2x_2}{\log^3(x/y)}\right].$$

We have, for n lying in such an interval $(x_1, x_2]$,

$$\tau(n, \tfrac{x_2}{z}, \tfrac{x_1}{y}) \ge 1 \implies \tau(n, y, z) \ge 1 \implies \tau(n, \tfrac{x_1}{z}, \tfrac{x_2}{y}) \ge 1.$$

We obtain the upper bound

$$H(x, y, z) \le \frac{x}{\log^2(x/y)} + \sum_{x_1, x_2} H(x_2, \tfrac{x_1}{z}, \tfrac{x_2}{y}) - H(x_1, \tfrac{x_1}{z}, \tfrac{x_2}{y}).$$

For large enough y_0,

$$\log\left(\frac{x_2}{y}\right) \ge \log\left(\frac{x}{y\log^2(x/y)}\right) \ge \frac{1}{2}\log\left(\frac{x}{y}\right),$$

and thus

$$x_2 - x_1 \ge \frac{x_2}{\log^4(x_2/y)}.$$

Also,

$$\log\left(\frac{x_2/y}{x_1/z}\right) \asymp \eta, \qquad \frac{x_1}{z} \le \frac{x_1}{\sqrt{x}} \le \sqrt{x_2},$$

so that by Theorem 2 and the part of Corollary 1 already proved,

$$H(x_2, \tfrac{x_1}{z}, \tfrac{x_2}{y}) - H(x_1, \tfrac{x_1}{z}, \tfrac{x_2}{y}) \ll \frac{x_2 - x_1}{x_2} H(x_2, \tfrac{x_1}{z}, \tfrac{x_2}{y}) \ll \frac{x_2 - x_1}{x} H(x, \tfrac{x}{z}, \tfrac{x}{y}).$$

Summing over intervals $(x_1, x_2]$ gives the desired upper bound. The lower estimate is obtained in the same way starting from

$$H(x, y, z) \ge \sum_{x_1, x_2} H(x_2, \tfrac{x_2}{z}, \tfrac{x_1}{y}) - H(x_1, \tfrac{x_2}{z}, \tfrac{x_1}{y}).$$

This completes the proof of Theorems 1, 2, and 3.

Proof of Theorems 4 *and* 5 *when* $z \le y^C$. Suppose $y_0(r)+1 \le y+1 \le z \le x^{5/8}$. When $z \le e^{100rC}y$, the trivial bound $H_r(x, y, z) \le H(x, y, z)$ suffices for an upper bound. For $e^{100rC}y \le z \le y^C$, the desired upper bound follows from Lemmas 3.4 and 3.9, plus the lower bound for $H(x, y, z)$ given in Theorem 1.

If $r = 1$ and $\eta \le \frac{1}{\log^2 y}$, arguing as in the lower bounds for $H(x, y, z)$, we have

$$H_1(x, y, z) \ge \sum_{y<d\le z} \left\lfloor \frac{x}{d} \right\rfloor - O(\eta^2 x \log y) \gg \eta x \gg H(x, y, z).$$

If $r \ge 1$ and $\frac{1}{\log^2 y} \le \eta \le \frac{1}{100}$, combining (5.2) (taking $\alpha = \frac{c}{300r}$ and $g = r$) and Lemma 4.3 with $c' = c/3$ gives

$$H_r(x, y, z) \gg_{r,c} \frac{\beta\eta^r x}{(\log y)^{r+1}} \sum_{(1+\frac{\beta}{100})v \le k \le \min(1+\beta, \log 4)v} \frac{(2^r v \log 2)^k}{k!}. \tag{5.7}$$

When $r = 1$, Lemma 2.6 gives

$$H_1(x, y, z) \gg_c \frac{\beta x}{\max(1, -\xi)(\log y)^{G(\beta)}}.$$

When $r \geq 2$, $z \geq z_0(y)$ and $\eta \leq \frac{1}{100}$, the sum on k in (5.7) is dominated by the term $k = \lfloor (1+\beta)v \rfloor$, whence by Stirling's formula, (1.3) and Theorem 1,

$$\begin{aligned} H_r(x,y,z) &\gg_{r,c} \frac{\beta x}{(\log\log y)^{1/2}(\log y)^{G(\beta)}} \\ &\gg_{r,c} H(x,y,z)\frac{\max(1,-\xi)}{\sqrt{\log\log y}}. \end{aligned}$$

Next, let $\alpha = \frac{c}{300r}$, $\gamma = 2^{-2M-1/\alpha}$ and suppose $\frac{1}{100} \leq \eta \leq \gamma \log y$ and $c' = c/3$. When $\eta \geq 1000r \cdot 3^{2r}$, apply Lemma 4.4 and the $g=1$ case of (5.5). Otherwise apply Lemma 4.3 and the $g=r$ case of (5.5). In either case, we obtain

$$\begin{aligned} H_r(x,y,z) &\gg_{r,c} x\frac{u^\delta(\log(2+\eta))^{\nu(r)+1}}{\eta(-\log u)^{3/2}} \\ &\gg_{r,c} H(x,y,z)\frac{(\log(2+\eta))^{\nu(r)+1}}{\eta}. \end{aligned} \tag{5.8}$$

Note that by (1.3), $G(\beta) = \delta + O(1/\log\log y)$ for $1.01y \leq z \leq 2y$, and so $(\log y)^{G(\beta)} \asymp (\frac{\log y}{\eta})^\delta$ in this range. The desired lower bound for $H_r(x,y,z)$ now follows from (5.8) and the upper bound for $H(x,y,z)$ in Theorem 1. When $y^{1+\gamma} \leq z \leq \min(y^C, x^{1/2-c/3})$ (for $r=1$, take $C=10$), we take the $h=1$ term in the sum in Lemma 4.4 (i), obtaining

$$H_r(x,y,z) \gg_{r,c,C} \frac{x(\log\log y)^{\nu(r)+1}}{\log y}. \tag{5.9}$$

Finally, when $x^{1/2-c/3} < z \leq x^{5/8}$, $z \leq y^C$ and $yz \leq x^{1-c}$, the desired lower bound is given in Lemma 4.4 (ii) since $y \leq x^{1/2-2c/3}$ and thus $\min(y, z/y) \geq y^{c/3}$.

Proof of Theorem 5 (1.7). Apply Lemma 4.4 (ii) with $c = \frac{1}{16}$.

Proof of Theorem 4 *when* $y^{10} \leq z \leq x^{5/8}$. This proof is quite simple and does not depend on the results of Sections 3 and 4. In this range, $H(x,y,z) \gg x$. Write each n with $\tau(n,y,z) = 1$ in the form

$$n = klm, \qquad P^+(k) \leq y, \qquad l \in \mathscr{P}(y,z), \qquad P^-(m) > z.$$

If $p^2 | l$ for some prime p, then $p > \sqrt{z}$ and thus the number of such n is $\ll x/\sqrt{z}$. Otherwise, $l=1$ or l is prime. Also $k \leq y^2$, for otherwise k has at least 2 divisors in $(y,z]$. Thus, $kl \leq y^2 z \leq z^{3/2} \leq x^{15/16}$.

The number of n with $m=1$ is $\leq x^{15/16}$. Now suppose $m > 1$. For each k, l, $x/kl \geq x^{1/16} \geq z^{1/16}$. By Lemma 2.3, the number of m is $\ll x/(kl\log z)$. Clearly k and l can't both be 1. If $k=1$, then l is prime and by Lemma 2.2, the number of such n is

$$\ll \frac{x}{\log z}\sum_{y<p\leq z}\frac{1}{p} \ll \frac{x(\log\log z - \log\log y)}{\log z}.$$

If $l > 1$ and $k > 1$, then $k \le y$ and also $z/P^-(k) < l \le z$. For a given k, the sum over l of $1/l$ is $\ll \frac{\log P^-(k)}{\log z}$. Also,

$$\sum_{2 \le k \le y} \frac{\log P^-(k)}{k} \le \sum_{p \le y} \frac{\log p}{p} \sum_{k' \in \mathscr{P}(p-1/2,y)} \frac{1}{k'} \ll (\log y) \log\log y.$$

Thus, the number of such n is

$$\ll \frac{x (\log y) \log\log y}{\log^2 z} \ll \frac{x \log\log y}{\log z}.$$

The last case to consider is $l = 1$ and $k > 1$. Here $k/P^-(k) \le y < k$. and we write $k = pk'$, where $p = P^-(k)$, $P^-(k') \ge p$ and $y/p < k' \le y$. By Lemma 2.3 and partial summation, the number of such n is

$$\ll \frac{x}{\log z} \sum_{p \le y} \frac{1}{p} \sum_{\substack{P^-(k') \ge p \\ y/p < k' \le y}} \frac{1}{k'} \ll \frac{x}{\log z} \sum_{p \le y} \frac{1}{p} \ll \frac{x \log\log y}{\log z}.$$

Putting these estimates together proves the upper bound.

For the lower bound, first note that if $kl \le \frac{x}{2z}$, then by Lemma 2.3, the number of m is $\gg \frac{x}{kl \log z}$. The number of n with $k = 1$, l a prime in $(y, z^{1/3})$ and $P^-(m) > z$ is

$$\gg \frac{x}{\log z} \sum_{y < p \le z^{1/3}} \frac{1}{p} \gg \frac{x(\log\log z - \log\log y)}{\log z}.$$

Next, let $l = 1$ and put $k = ph$, where $10 < p \le y^{1/4}$, $P^-(h) > p$ and $y/p < h \le y$. The number of such n is, by Lemma 2.3 and partial summation,

$$\begin{aligned} &\gg \frac{x}{\log z} \sum_{10 < p \le y^{1/4}} \frac{1}{p} \sum_{\substack{j \in \mathbb{Z} \\ y/p < 2^j \le y/2}} \frac{\Phi(2^{j+1}, p) - \Phi(2^j, p)}{2^j} \\ &\gg \frac{x}{\log z} \sum_{10 < p \le y^{1/4}} \frac{1}{p} \gg \frac{x \log\log y}{\log z}. \end{aligned}$$

This completes the lower bound.

Remarks. By extending the methods used to prove Theorem 4 when $y^{10} \le z \le x^{5/8}$, it should be possible to determine the order of $H_r(x, y, z)$ for $y^{10} \le z \le x^{1/2}$ for any fixed $r \ge 2$. We conjecture that for each $r \ge 1$ and $y^{10} \le z \le x^{1/2}$,

$$\frac{H_r(x, y, z)}{x} \asymp \frac{Q_r(\log\log y, \log\log z)}{\log z}$$

for some polynomial Q_r.

6. Initial sums over $L(a;\sigma)$ and $L_s(a;\sigma)$

The object of this section is to prove Lemmas 3.2 and 4.1. The upper bounds are more complicated, due to having to count integers with $d|n$, $y < d \le z$ and $P^+(d) = z^{o(1)}$. It is convenient to work with divisors $d|n$ with $P^+(d) < P^+(n)$. If this is not the case, then the complementary divisor n/d does satisfy $P^+(n/d) < P^+(n)$ (here we assume that $P^+(n)^2 \nmid n$, the number of exceptions being very small). We put n into a very short interval, so that when $y < d \le z$, n/d lies in an interval $(y', z']$ with $\log(z'/y') \approx \log(z/y)$.

LEMMA 6.1. *Suppose y, z, x_1, x_2 are positive real numbers satisfying*

$$100 \le y < z = e^\sigma y \le x_1^{3/4}, \quad z \le y^{\log\log y},$$

and

$$\sigma \ge \frac{1}{10\log^2 z}, \quad \frac{3^{10}x_1}{\log^{10} z} \le x_2 - x_1 \le \frac{x_1}{\log^4 z}.$$

Then

$$H(x_2, y, z) - H(x_1, y, z) \ll (x_2 - x_1)\big[S(y;\sigma) + S(x_2/z;\sigma)\big].$$

If in addition $r \ge 1$ and $y \ge y_0(r)$, then

$$H_r(x_2, y, z) - H_r(x_1, y, z) \ll_r (x_2 - x_1) \sum_{\substack{1\le s\le r\\ \nu(s)\le\nu(r)}} \big[S_s(y;\sigma) + S_s(x_2/z;\sigma)\big].$$

Proof. Let $\mathscr{A}$ be the set of integers $n \in (x_1, x_2]$ satisfying

(i) $\tau(n, y, z) \ge 1$;

(ii) $\tau(n, x_1/z, x_2/z) = 0$;

(iii) if p is prime, $p|n$ and $p > \log^{10} z$, then $p^2 \nmid n$;

(iv) if $d|n$ with $y < d \le z$, then $P^+(d) > \log^{20} z$ and $P^+(n/d) > \log^{20}(x_2/y)$.

Let $\mathscr{A}_r$ be the set of $n \in \mathscr{A}$ with $\tau(n, y, z) = r$. Since $x_2 - x_1 \ge x_1/z$, the number of integers in $(x_1, x_2]$ not satisfying (ii) is at most

$$\sum_{x_1/z\le d\le x_2/z} \left(\frac{x_2 - x_1}{d} + 1\right) \ll \frac{(x_2 - x_1)^2}{x_1} \ll \frac{x_2 - x_1}{\log^4 z}.$$

The number of integers in $(x_1, x_2]$ failing (iii) is

$$\le \sum_{\log^{10} z<p\le\sqrt{x_2}} \left(\frac{x_2 - x_1}{p^2} + 1\right) \ll \frac{x_2 - x_1}{\log^{10} z}.$$

Put $y_1 = y$, $z_1 = z$, $y_2 = x_2/z$ and $z_2 = x_2/y$. Every $n \in (x_1, x_2]$ satisfying (i) and (ii) can be written in the form

$$n = m_1 m_2, \qquad y_j < m_j \le z_j \ (j = 1, 2). \tag{6.1}$$

Note that $z_2 \ge x_1^{1/4} \ge z_1^{1/3}$ and so $y_j \ge z_j^{1/(3\log\log z_j)}$ for $j = 1, 2$. By Lemma 2.5, the number of integers $\le x_2$ not satisfying (iv) is

$$\le \sum_{j=1}^{2} \sum_{\substack{m_j > y_j \\ P^+(m_j) \le \log^{20} z_j}} \frac{x_2}{m_j} \ll x_2 \sum_{j=1}^{2} (\log\log z_j) e^{-\frac{\log y_j}{80\log\log z_j}}$$
$$\ll \frac{x_2}{\log^{100} z} \ll \frac{x_2 - x_1}{\log^{10} z}.$$

We conclude that

$$\begin{aligned} H(x_2, y, z) - H(x_1, y, z) &\le |\mathscr{A}| + O\left(\frac{x_2 - x_1}{\log^4 z}\right), \\ H_r(x_2, y, z) - H_r(x_1, y, z) &\le |\mathscr{A}_r| + O\left(\frac{x_2 - x_1}{\log^4 z}\right). \end{aligned} \tag{6.2}$$

We follow similar procedures for bounding $|\mathscr{A}|$ and $|\mathscr{A}_r|$. First, writing $n \in \mathscr{A}$ in the form (6.1), we see by (iv) that $P^+(m_j) > \log^{20} z_j$ for $j = 1, 2$. By (iii), $P^+(m_1) \ne P^+(m_2)$. Suppose that

$$p = P^+(m_j) < P^+(m_{3-j}) \qquad (j = 1 \text{ or } j = 2), \tag{6.3}$$

and write

$$n = abp, \quad P^+(a) < p < P^-(b), \quad b > p. \tag{6.4}$$

We have $\tau(a, y_j/p, z_j/p) \ge 1$, which implies

$$\log(y_j/p) \in \mathscr{L}(a; \sigma).$$

Each $n \in \mathscr{A}_r$ may be written uniquely as

$$n = m_{11}m_{12} = m_{21}m_{22} = \cdots = m_{r1}m_{r2},$$

where $y_1 < m_{11} < \cdots < m_{r1} \le z_1$ and $y_2 < m_{r2} < \cdots < m_{12} \le z_2$. There may be divisors of n lying in $(y_1, z_1] \cap (y_2, z_2]$, in which case $m_{i1} = m_{i'2}$ for some pairs i, i'. Let $p_{ij} = P^+(m_{ij})$ for each i, j. Since $p_{i1} \ne p_{i2}$ for each i, $p_{ij} \ne P^+(n)$ for exactly r pairs of indices i, j. Therefore, there is an integer s with $1 \le s \le r$ and $\nu(s) \le \nu(r)$, a $j \in \{1, 2\}$ and a prime $p < P^+(n)$, so that exactly s of the primes $p_{1j}, \ldots, p_{rj}$ are equal to p. Writing n in the form (6.4), we see that $\tau(a, y_j/p, z_j/p) = s$, and thus

$$\log(y_j/p) \in \mathscr{L}_s(a; \sigma).$$

By (iv) and (6.4),

$$b > p > \log^{20} z_j \ge \log^{20}(z^{1/3}) \tag{6.5}$$

and so

$$\frac{x_2 - x_1}{ap} \ge \frac{x_2 - x_1}{x_2} p \ge \frac{p}{\log^{10}(z^{1/3})} > p^{1/2}.$$

By Lemma 2.3, given a and p, the number of possible b is at most

$$\Phi\left(\frac{x_2}{ap}, p\right) - \Phi\left(\frac{x_1}{ap}, p\right) \ll \frac{x_2 - x_1}{ap \log p}.$$

Let $Q_j(a) := \max(\log^{20} z_j, P^+(a))$, so that $p > Q_j(a)$. By (6.4) and (6.5),

$$(6.6)\qquad \begin{aligned} |\mathscr{A}| &\ll (x_2 - x_1) \sum_{j=1}^{2} \sum_{P^+(a) \le z_j} \frac{1}{a} \sum_{\substack{\log(y_j/p) \in \mathscr{L}(a;\sigma) \\ p > Q_j(a)}} \frac{1}{p \log p}, \\ |\mathscr{A}_r| &\ll (x_2 - x_1) \sum_{j=1}^{2} \sum_{\substack{1 \le s \le r \\ \nu(s) \le \nu(r)}} \sum_{P^+(a) \le z_j} \frac{1}{a} \sum_{\substack{\log(y_j/p) \in \mathscr{L}_s(a;\sigma) \\ p > Q_j(a)}} \frac{1}{p \log p}. \end{aligned}$$

Let $\mathscr{L} = \mathscr{L}(a;\sigma)$ or $\mathscr{L} = \mathscr{L}_s(a;\sigma)$, as appropriate. Then

$$\log(y_j/p) \in \mathscr{S} := \mathscr{L} \cap \left[-\sigma, \log\left(\frac{y_j}{Q_j(a)}\right)\right).$$

The set $\mathscr{L}$ is the disjoint union of intervals, each with a left endpoint $-\sigma + \log d$ for some $d|a$ or a right endpoint $\log d$ for some $d|a$. Therefore, $\mathscr{S}$ has the same property with the possible exception of one interval whose right endpoint is $\log(y_j/Q_j(a))$. Breaking long intervals into many short ones, we may partition $\mathscr{S}$ into intervals I_i, each of length $1/\log^{10} z_j$, and intervals I_i', each with left endpoint $-\sigma + \log d$ or right endpoint $\log d$ for some $d|a$ (with one possible exception) and of length $< 1/\log^{10} z_j$. If $I_i = [A, B)$, then $y_j e^{-B} \ge \log^{20} z_j$. By Lemma 2.1,

$$\sum_{\log(y_j/p) \in I_i} \frac{1}{p \log p} \le \frac{\pi(y_j e^{-A}) - \pi(y_j e^{-B})}{y_j e^{-B} \log(y_j e^{-B})} \ll \frac{B - A}{\log^2(y_j e^{-B})}.$$

Since $B \le \log a$ and $B \le \log(y_j/P^+(a))$ when $a > 1$, we have

$$\log(y_j e^{-B}) \ge \log \max(y_j/a, P^+(a)).$$

Adding the contributions of all intervals I_i gives

$$\sum_{\log(y_j/p) \in \cup I_i} \frac{1}{p \log p} \ll \frac{\operatorname{meas}(\mathscr{L})}{\log^2(y_j/a + P^+(a))}.$$

Trivially, for each i

$$\sum_{\log(y_j/p) \in I_i'} \frac{1}{p \log p} \le \sum_{\log(y_j/m) \in I_i'} \frac{1}{m} \ll \frac{1}{\log^{10} z_j}.$$

The number of intervals I_i' is $\le 2\tau(a) + 1 \le 3\tau(a)$, and thus

$$\sum_{\log(y_j/p) \in \mathscr{S}} \frac{1}{p \log p} \ll \frac{\operatorname{meas}(\mathscr{L})}{\log^2(y_j/a + P^+(a))} + \frac{\tau(a)}{\log^{10} z_j}.$$

Next, we sum on a, j and s in (6.6) and use

$$\sum_{P^+(a)\le z_j} \frac{\tau(a)}{a} = \prod_{p\le z_j}\left(1+\frac{2}{p}+\frac{3}{p^2}+\cdots\right) \ll \log^2 z_j.$$

By (6.2), this gives

$$H(x_2,y,z)-H(x_1,y,z) \ll (x_2-x_1)\left[\frac{1}{\log^4 z}+S(y_1;\sigma)+S(y_2;\sigma)\right],$$
$$H_r(x_2,y,z)-H_r(x_1,y,z) \ll (x_2-x_1)\left[\frac{1}{\log^4 z}+\sum_{\substack{1\le s\le r\\ \nu(s)\le\nu(r)}}\big(S_s(y_1;\sigma)+S_s(y_2;\sigma)\big)\right].$$

Since $L_1(1;\sigma)=\sigma$, we have

$$S(y_1;\sigma)\ge S_1(y_1;\sigma)\ge \frac{\sigma}{\log^2 y_1}\ge \frac{1}{10\log^4 z},$$

and this completes the proof. □

Proof of Lemma 3.2. Let $x_0=\max(x-\Delta, x/\log^{100} z)$, partition $(x_0,x]$ into intervals $(x_1,x_2]$ with $x_2-x_1\asymp x_1/\log^9 z$ and apply Lemma 6.1 (with $\sigma=\eta$) to each. For each pair (x_1,x_2), we have $z\le x^{5/8}\le x_1^{3/4}$, $\sigma\ge\frac{1}{\log y}\ge\frac{1}{\log^2 z}$ and $x_2/z\ge x^{1/3}\ge y^{1/2}$. Thus

$$H(x,y,z)-H(x-\Delta,y,z)\ll \frac{x}{\log^{100} z}+\Delta\max_{y^{1/2}\le t\le x} S(t;\eta),$$
$$H_r(x,y,z)\ll \frac{x}{\log^{100} z}+x\max_{\substack{1\le s\le r\\ \nu(s)\le\nu(r)}}\ \max_{y^{1/2}\le t\le x} S_s(t;\eta).$$

Since $L_1(1;\eta)=\eta\ge\frac{1}{\log y}$, we have $S(y;\eta)\ge S_1(y;\eta)\ge 1/\log^3 y$. The lemma follows. □

Turning now to the lower bounds, our task is simpler because we may restrict the integers n in any manner we choose.

Proof of Lemma 4.1. First we derive the lower bound for $H^*(x,y,z)-H^*(x-\Delta,y,z)$. Consider all square-free integers $n=abp\in(x-\Delta,x]$, where

(i) $a\le y^{1/8}$;

(ii) $q|b$ and q prime implies $y^{1/4}<q<y^{7/8}$ or $q>z$;

(iii) p is prime and $\log(y/p)\in\mathscr{L}(a;\eta)$.

Condition (iii) ensures that ap has a divisor in $(y,z]$, and so does n as well. By (i) and (iii), $y/z\le y/p\le a\le y^{1/8}$; thus $y^{7/8}\le p\le z$. Hence each n has a unique representation in the form abp. For fixed a and p,

$$\frac{x}{ap}\ge\frac{x}{y^{1/8}z}\ge y^{1/2}.$$

Also, $\Delta \ge x/\log^{10} z$, so that

$$\frac{\Delta}{ap} \ge \frac{x}{ap\log^{10}(x/ap)}.$$

If $y^{1/2} \le x/(ap) \le 2z$, we count b with $P^-(b) > y^{1/4}$ (so automatically $P^+(b) < y^{7/8}$). By Lemma 2.3, the number of such b is $\gg \Delta/(ap\log y)$. When $x/(ap) > 2z$, take b so that $P^-(b) > z$. By Lemma 2.3, the number of such b is $\gg \frac{\Delta}{ap\log z} \gg \frac{\Delta}{ap\log y}$ as well. The number of b which are not square-free is at most

$$\sum_{q>y^{1/4}} \frac{x}{apq^2} \ll \frac{x}{apy^{1/4}} \ll \frac{\Delta}{apy^{1/5}}.$$

Thus, we obtain

$$H^*(x,y,z) - H^*(x-\Delta,y,z) \gg \frac{\Delta}{\log y} \sum_{\substack{a\le y^{1/8}\\ \mu^2(a)=1}} \frac{1}{a} \sum_{\log(y/p)\in\mathscr{L}(a;\eta)} \frac{1}{p}.$$

For each a satisfying (i), $\mathscr{L}(a;\eta)$ consists of a disjoint union of intervals, each with length $\ge \eta$. Breaking long intervals into shorter ones, we may partition $\mathscr{L}(a;\eta)$ into intervals each with length between $\frac{1}{2\log^2 y}$ and $\frac{1}{\log^2 y}$. If $[v, v+w]$ is one such interval, then $v+w \le \log a \le \frac{\log y}{8}$ and thus by Lemma 2.2,

$$\sum_{v\le\log(y/p)<v+w} \frac{1}{p} = \sum_{ye^{-v-w}<p\le ye^{-v}} \frac{1}{p} \gg \frac{w}{\log y}.$$

Combining all such intervals gives

$$\sum_{\log(y/p)\in\mathscr{L}(a;\eta)} \frac{1}{p} \gg \frac{L(a;\eta)}{\log y},$$

which completes the proof of the first assertion.

Next we derive the lower bound for $H_r(x,y,z)$. Consider all integers $n = abp \le x$, where

(i) $a \le y^{2c}$;

(ii) $\log(y/p) \in \mathscr{L}_r(a;\eta)$;

(iii) $P^-(b) > z$.

We have $y^{1-2c} \le p \le z$, and so each n has a unique representation as abp. If $d|n$ with $y < d \le z$, then $d = pd'$ for some $d'|a$, thus (ii) ensures that $\tau(n,y,z) = r$. Since $ap \le z^{1+2c} \le x/4z$, for a given pair a,p, Lemma 2.3 implies that the number of b is $\gg \frac{x}{ap\log z} \gg \frac{x}{Cap\log y}$. In contrast with $\mathscr{L}(a;\eta)$, $\mathscr{L}_r(a;\eta)$ may not consist only of intervals of length $\ge \eta$. With a fixed, partition $\mathscr{L}_r(a;\eta)$ as

$$\mathscr{L}_r(a;\eta) = \left(\bigcup_{i=1}^N I_i\right) \cup \left(\bigcup_{j=1}^M I_j'\right),$$

where each I_i is an interval of length $w = 1/\log^{3r+7} y$, and each I'_j is an interval of length $< w$ and with a left endpoint $-\eta + \log d$ or a right endpoint $\log d$ for some $d|a$. Clearly

$$\sum_{j=1}^{M} \mathrm{meas}(I'_j) \le \frac{2\tau(a)}{\log^{3r+7} y}.$$

Consider one interval $I_i = [v, v+w]$. Since $v \ge -\eta$, $e^{v+w} \le a \le y^{2c}$ and $y \ge y_0(r,c)$, Lemma 2.2 gives

$$\sum_{\log(y/p)\in I_i} \frac{1}{p} \gg_C \frac{w}{\log y}.$$

Adding the contributions of all intervals I_i gives

$$\sum_{\log(y/p)\in \mathscr{L}_r(a;\eta)} \frac{1}{p} \ge K\left(\frac{L_r(a;\eta)}{\log y} - \frac{2\tau(a)}{\log^{3r+8} y}\right),$$

where K is a positive constant which depends on C. Summing over a gives

$$H_r(x,y,z) \gg_{r,c,C} \frac{x}{\log^2 y}\left(\sum_{a\le y^{2c}} \frac{L_r(a;\eta)}{a} - \frac{2}{\log^{3r+7} y}\sum_{a\le y^{2c}} \frac{\tau(a)}{a}\right).$$

The second sum on a is $O(\log^2 y)$, thus it suffices to show that the first sum on a is $\gg (\log y)^{-3r-2}$. By standard prime number estimates, there is an interval of length $100r\log\log y$ contained in $[\log^3 y, 2\log^3 y]$ which contains r primes $p_1 < \cdots < p_r$. If $a_0 = p_1\cdots p_r$ and y is large enough, then $\log p_r - \log p_1 \le \eta/10$ and $\log(p_1p_2) - \log(p_r) \ge \log\log y$. Hence

$$L_r(a_0;\eta) \ge \min(\eta, \log\log y) - \log(p_r/p_1) \ge \frac{1}{2\log^2 y}.$$

Therefore

$$\frac{L_r(a_0;\eta)}{a_0} \gg_r \frac{1}{(\log y)^{3r+2}},$$

and this completes the proof. $\square$

7. Upper bounds in terms of $S^*(t;\sigma)$

In this section we prove Lemmas 3.3 and 3.4.

Proof of Lemma 3.3. We will show that for $0 < \sigma \le \log t$,

$$S(t;\sigma) \ll (1+\sigma)\hat{S}(t;\sigma), \tag{7.1}$$

where

$$\hat{S}(t;\sigma) = \sum_{a\in\mathscr{P}(e^\sigma, te^\sigma)} \frac{L(a;\sigma)}{a\log^2(t^{7/8}/a + P^+(a))}, \tag{7.2}$$

and then prove

$$\hat{S}(t;\sigma) \ll S^*(t;\sigma) \qquad (t \text{ large}, \sigma > 0). \tag{7.3}$$

First, assume $\sigma \ge \log 2$, else (7.1) is trivial. In (3.3), write each $a = a_1 a_2$, where $P^+(a_1) \le e^\sigma < P^-(a_2)$. By Lemma 3.1 (iii) and (iv),

$$L(a;\sigma) \le L(a_2; \sigma + \log a_1) \le \left(1 + \frac{\log a_1}{\sigma}\right) L(a_2;\sigma). \tag{7.4}$$

The contribution to $S(t;\sigma)$ from those a with $a_1 \ge t^{1/8}$ is thus

$$\ll \frac{1}{\sigma} \sum_{\substack{P^+(a_1)\le e^\sigma \\ a_1 > t^{1/8}}} \frac{\log a_1}{a_1 \log^2 P^+(a_1)} \sum_{a_2 \in \mathscr{P}(e^\sigma, te^\sigma)} \frac{L(a_2;\sigma)}{a_2}. \tag{7.5}$$

By Lemma 2.5, the sum on a_1 in (7.5) is

$$\begin{aligned}
&\le 4 \sum_{j\in\mathbb{Z}: t^{2^{-j}} \le e^{2\sigma}} \frac{2^{2j}}{\log^2 t} \sum_{\substack{a_1 > t^{1/8} \\ t^{2^{-j-1}} < P^+(a_1) \le t^{2^{-j}}}} \frac{\log a_1}{a_1} \\
&\ll \sum_{j\in\mathbb{Z}: 2^j \ge \frac{\log t}{2\sigma}} 2^j e^{-2^{j-5}} \\
&\ll \frac{\sigma^2}{\log^2 t}\left[\left(\frac{\log t}{\sigma}\right)^3 \exp\left\{-\frac{1}{64}\frac{\log t}{\sigma}\right\}\right] \ll \frac{\sigma^2}{\log^2 t}.
\end{aligned}$$

It follows that the expression in (7.5) is $\ll \sigma \hat{S}(t;\sigma)$. By (7.4), $P^+(a) \ge P^+(a_2)$ and Lemma 2.5, the contribution to $S(t;\sigma)$ from those a with $a_1 \le t^{1/8}$ is at most

$$\sum_{P^+(a_1)\le e^\sigma} \frac{1 + \frac{1}{\sigma}\log a_1}{a_1} \sum_{a_2\in\mathscr{P}(e^\sigma, te^\sigma)} \frac{L(a_2;\sigma)}{a_2 \log^2(t^{7/8}/a_2 + P^+(a_2))} \ll \sigma \hat{S}(t;\sigma).$$

This proves (7.1). Next, in (7.2) write $a = a_1 a_2$ with

$$a_1 = \prod_{p^\beta \| a, \beta \ge 2} p^\beta,$$

so that $(a_1, a_2) = 1$ and a_2 is square-free. By Lemma 3.1,

$$L(a;\sigma) \le \tau(a_1) L(a_2;\sigma) \le \sigma \tau(a_1)\tau(a_2). \tag{7.6}$$

Also

$$\sum_{a_1 > w} \frac{\tau(a_1)}{a_1} \ll \sum_{a_1 > w} a_1^{-3/4} \ll w^{-1/4} \quad (w \ge 1). \tag{7.7}$$

The contribution to $\hat{S}(t;\sigma)$ coming from those a with $a_1 \ge t^{1/8}$ is thus

$$\ll \sigma \sum_{a_2 \in \mathscr{P}^*(e^\sigma, te^\sigma)} \frac{\tau(a_2)}{a_2} \sum_{a_1 \ge t^{1/8}} \frac{\tau(a_1)}{a_1} \ll \frac{\sigma \log^2 t}{t^{1/32}} \ll S^*(t;\sigma), \tag{7.8}$$

where we used (3.6) in the last step. Using $P^+(a) \ge P^+(a_2)$ and (7.6), we see that the contribution to $\hat{S}(t;\sigma)$ from those a with $a_1 < t^{1/8}$ is

$$\le \sum_{a_1} \frac{\tau(a_1)}{a_1} \sum_{a_2 \in \mathscr{P}^*(e^\sigma, te^\sigma)} \frac{L(a_2;\sigma)}{a_2 \log^2(t^{3/4}/a_2 + P^+(a_2))} \ll S^*(t;\sigma).$$

This proves (7.3). □

Lemma 3.4 depends on the distribution of the first $r+1$ divisors of typical integers. Throughout the remainder of this section, let $d_j(n)$ denote the j-th smallest divisor of n.

Lemma 7.1. *Let n have prime factorization $n = p_1^{e_1} \cdots p_f^{e_f}$, where $p_1 < \cdots < p_f$. Let $N = \tau(n)$, $1 \le \ell \le N-1$ and define v uniquely by*

$$(e_1+1)\cdots(e_{v-1}+1)|\ell, \qquad (e_1+1)\cdots(e_v+1) \nmid \ell. \tag{7.9}$$

Then $d_{\ell+1}(n)/d_\ell(n) \le p_v$.

Remarks. Lemma 7.1 is nearly best possible, e.g. if one takes $e_1 = \cdots = e_f = 1$ and chooses the prime divisors of n so that $p_i > p_{i-1}^4$ for $2 \le i \le f$, then $v = \nu(\ell)+1$ and

$$\frac{d_{\ell+1}(n)}{d_\ell(n)} = \frac{p_v}{p_1 \cdots p_{v-1}} \ge p_v^{2/3}$$

for every ℓ. Moreover, this represents the typical case and is the origin of the exponent $\nu(r)+1$ appearing in Lemma 3.4 and Theorem 5.

Proof. We apply induction on f, the case $f=1$ being trivial. Assume the statement is true for $f=m$ and take an integer n with prime factorization $n = p_1^{e_1}\cdots p_{m+1}^{e_{m+1}}$, where $p_1 < \cdots < p_{m+1}$. Put $N=\tau(n)$, suppose $1 \le \ell \le N-1$ and define v by (7.9). The conclusion is trivial if $v = m+1$; so suppose $v \le m$. Let $h = p_1^{e_1}\cdots p_m^{e_m}$ and $E = e_{m+1}+1$. For $0 \le j \le E-1$, let

$$a_j = |\{i \le \ell : p_{m+1}^j \| d_i(n)\}| = |\{d|h : dp_{m+1}^j \le d_\ell(n)\}|.$$

Since $a_0 + \cdots + a_{E-1} = \ell$, by (7.9) there is a j so that $(e_1+1)\cdots(e_v+1) \nmid a_j$ (in particular $1 \le a_j \le \tau(h)-1$). Define u by

$$(e_1+1)\cdots(e_{u-1}+1)|a_j, \qquad (e_1+1)\cdots(e_u+1) \nmid a_j,$$

so that $1 \le u \le v$. Finally, $d_{a_j+1}(h)p_{m+1}^j$ is a divisor of n that is larger than $d_\ell(n)$. Therefore, by the induction hypothesis,

$$\frac{d_{\ell+1}(n)}{d_\ell(n)} \le \frac{d_{a_j+1}(h)p_{m+1}^j}{d_{a_j}(h)p_{m+1}^j} \le p_u \le p_v. \qquad \square$$

Using Lemma 7.1, we can obtain upper bounds of averages of ratios of consecutive divisors of numbers.

LEMMA 7.2. *Fix $\ell \ge 1$. Uniformly in $x \ge 1$, $y \ge 4$,*

$$\sum_{\substack{P^+(n)\le y\\ n\ge x\\ \tau(n)\ge \ell+1}} \frac{\log(d_{\ell+1}(n)/d_\ell(n))}{n} \ll_\ell (\log y)(\log\log y)^{\nu(\ell)+1} \exp\left\{-\frac{\log x}{4\log y}\right\}.$$

Proof. Denote by S the sum in the lemma. Let V be the largest number so that ℓ is the product of V integers, each at least 2 (if $\ell = 1$, set $V = 0$). Suppose $n > x$, $\tau(n) \ge \ell + 1$ and $P^+(n) \le y$. For each such n there is a unique v between 1 and $V + 1$, so that $n = p_1^{e_1} \cdots p_v^{e_v} m$, where $p_1 < \cdots < p_v$, $m \in \mathscr{P}(p_v, y)$ and

$$(e_1+1)\cdots(e_{v-1}+1)|\ell, \qquad (e_1+1)\cdots(e_v+1)\nmid \ell. \tag{7.10}$$

By Lemma 7.1, $d_{\ell+1}(n)/d_\ell(n) \le p_v$. Since $e_1 + \cdots + e_{v-1} \le \ell$, $m > x/y^{\ell+e_v}$. Thus $S \le AB$, where (ignoring the second condition in (7.10) and the condition $p_v > p_{v-1}$)

$$A = \sum_{v=1}^{V+1} \sum_{e_1,\dots,e_{v-1}} \sum_{p_1<\cdots<p_{v-1}\le y} \frac{1}{p_1^{e_1}\cdots p_{v-1}^{e_{v-1}}},$$
$$B = \sum_{\substack{p\le y\\ f\ge 1}} \frac{\log p}{p^f} \sum_{\substack{m\in\mathscr{P}(p,y)\\ m>x/y^{\ell+f}}} \frac{1}{m}.$$

For each tuple $v, e_1, \dots, e_{v-1}$ satisfying the first condition of (7.10), at most $\nu(\ell)$ of the numbers e_i can equal 1. Thus, by Lemma 2.2,

$$A \ll_v (\log\log y)^{\nu(\ell)}.$$

By Lemmas 2.2 and 2.5,

$$\begin{aligned} B &\ll (\log y) \sum_{\substack{p\le y\\ f\ge 1}} \frac{1}{p^f} \exp\left\{-\frac{\log(x/y^{\ell+f})}{4\log y}\right\} \\ &\le e^{\frac{\ell}{4}-\frac{\log x}{4\log y}} (\log y) \sum_{f=1}^{\infty} e^{f/4} \sum_{p\le y} p^{-f} \\ &\ll_\ell (\log y)(\log\log y) e^{-\frac{\log x}{4\log y}}. \end{aligned}$$

□

Proof of Lemma 3.4. In light of Lemma 3.2, it suffices to show that

$$S_s(t;\eta) \ll_r (\log \eta)^{\nu(r)+1} S^*(t;\eta) \tag{7.11}$$

for each s satisfying $1 \le s \le r$, $\nu(s) \le \nu(r)$ and each $t \ge y^{1/2}$. Let $\theta = \frac{\eta}{100Cr}$. By hypothesis, $\theta \ge 1$. In (3.4), write each $a = gh$, where $P^+(g) \le e^\theta < P^-(h)$. Put $m = \tau(g)$. We have $S_s(t,\eta) = T_1 + T_2 + T_3$, where T_1 is the sum over numbers a with $m \le s$ and $\nu(m) \le \nu(s)$; T_2 is the sum over those a with $g \le t^{1/16}$ and with $m > s$ or $\nu(m) > \nu(s)$; and T_3 is the sum over the remaining a.

For T_1, we use $L_s(gh;\eta) \le L(gh;\eta) \le mL(h;\eta)$, which is a consequence of Lemma 3.1. Also,

$$g \le e^{m\theta} \le e^{r\theta} \le (z/y)^{1/100C} \le y^{1/100} \le t^{1/50};$$

thus $t/(gh) + P^+(gh) \ge t^{15/16}/h + P^+(h)$. This gives

$$T_1 \le r \sum_{\substack{P^+(g)\le e^\theta \\ \tau(g)=m\le s, \nu(m)\le\nu(s)}} \frac{1}{g} \sum_{h\in\mathscr{P}(e^\theta, te^\eta)} \frac{L(h;\eta)}{h\log^2(t^{15/16}/h + P^+(h))}. \tag{7.12}$$

Factor each g as $p_1^{e_1}\cdots p_j^{e_j}$. Since $m = (e_1+1)\cdots(e_j+1)$ and $j \le \frac{\log s}{\log 2}$, there are at most $\nu(m)$ indices j with $e_j = 1$. Therefore, by Lemma 2.2, the sum on g in (7.12) is

$$\le \sum_j \sum_{\substack{e_1,\cdots,e_j \\ p_1,\dots,p_j\le e^\theta}} \frac{1}{p_1^{e_1}\cdots p_j^{e_j}} \ll_m (\log 2\theta)^{\nu(m)} \ll_r (\log\eta)^{\nu(r)}.$$

Consequently,

$$T_1 \ll_r (\log\eta)^{\nu(r)} \sum_{h\in\mathscr{P}(e^\theta, te^\eta)} \frac{L(h;\eta)}{h\log^2(t^{15/16}/h + P^+(h))}. \tag{7.13}$$

For g satisfying $m > s$ or $\nu(m) > \nu(s)$, we first prove that

$$L_s(gh;\eta) \le 4r\frac{L(h;\eta)}{\theta} \max_{\substack{1\le\ell\le\min(s,m-1) \\ \nu(\ell)\le\nu(s)}} \log\left(\frac{d_{\ell+1}(g)}{d_\ell(g)}\right). \tag{7.14}$$

Write $D_j = d_j(gh)$ for $1 \le j \le \tau(gh)$. If $\tau(gh, e^u, e^{u+\eta}) = s$, then $[u, u+\eta]$ contains an interval $[\alpha,\beta]$ so that $\tau(gh, e^\alpha, e^\beta) = 0$ and $\beta - \alpha \ge \eta/(s+1) \ge 2\theta$. Let

$$\begin{aligned} J &= \{1 \le j \le \tau(gh) - 1 : \log(D_{j+1}/D_j) \ge 2\theta\}, \\ M_j &= \{u \in \mathbb{R} : e^u < D_j \le e^{u+\eta}, \tau(gh, e^u, e^{u+\eta}) = s\} \quad (1 \le j \le \tau(gh)). \end{aligned}$$

If $\tau(gh, e^u, e^{u+\eta}) = s$, then either $u \in M_1 \cup M_{\tau(gh)}$ or for some $j \in J$, $u \in M_j \cup M_{j+1}$. Hence

$$L_s(a;\eta) \le (2|J| + 2)\max_j \operatorname{meas}(M_j). \tag{7.15}$$

Since $P^+(g) \le e^\theta$, if $j \in J$, then $g|D_j$ and $(g, D_{j+1}) = 1$, and thus D_j/g and D_{j+1} are consecutive divisors of h. Also $D_{j+1}/(D_j/g) \ge e^{2\theta}$, hence the intervals $(\log D_{j+1}-\theta, \log D_{j+1})$ are disjoint. All such intervals lie in $\mathscr{L}(h;\theta)$, thus $L(h;\theta)\ge\theta|J|$, and, since $L(h;\theta)\ge\theta$,

$$2|J|+2\le 2\frac{L(h;\theta)}{\theta}+2\le\frac{4L(h;\theta)}{\theta}. \tag{7.16}$$

Fix j and let

$$\mathscr{B}=\{b\in\mathbb{Z}:\max(1,j+1-s)\le b\le\min(j,\tau(gh)+1-s)\},$$

i.e., if $u\in M_j$, then the s divisors of gh lying in $(e^u,e^{u+\eta}]$ are $D_b,\dots,D_{b+s-1}$ for some $b\in\mathscr{B}$. Fix one such $b\in\mathscr{B}$ and let

$$M_{j,b}=\{u\in M_j: e^u<D_b<\cdots<D_{b+s-1}\le e^{u+\eta}\}.$$

For each $f|h$, let

$$\mathscr{K}_f=\{j:d_j(g)f\in\{D_b,\dots,D_{b+s-1}\}\},$$

which is a set of consecutive integers. Also, since $d_{j+1}(g)/d_j(g)\le e^\theta<e^{\frac{\eta}{s+1}}$, each set $\mathscr{K}_f$ contains 1 or m. Let $K_f=|\mathscr{K}_f|$. Since $\sum_f K_f=s$, there is some f so that $K_f\ge1$ and $\nu(K_f)\le\nu(s)$. We have $K_f<m$, otherwise $s\ge m$ and $\nu(K_f)=\nu(m)>\nu(s)$. If $\mathscr{K}_f=\{1,2,\dots,K_f\}$, then $D_{b+s}\le d_{K_f+1}(g)f$, so that

$$\operatorname{meas}(M_{j,b})\le\log\left(\frac{d_{K_f+1}(g)}{d_{K_f}(g)}\right).$$

Likewise, if $\mathscr{K}_f=\{m-K_f+1,\dots,m\}$, then

$$\operatorname{meas}(M_{j,b})\le\log\left(\frac{d_{m-K_f+1}(g)}{d_{m-K_f}(g)}\right)=\log\left(\frac{d_{K_f+1}(g)}{d_{K_f}(g)}\right),$$

by the symmetry of the divisors of g. Since $|\mathscr{B}|\le s$,

$$\operatorname{meas}(M_j)\le s\max_{\substack{1\le\ell\le\min(s,m-1)\\ \nu(\ell)\le\nu(s)}}\log\left(\frac{d_{\ell+1}(g)}{d_\ell(g)}\right).$$

Combined with (7.15) and (7.16), this proves (7.14).

If $g\le t^{1/16}$, then $t/gh+P^+(gh)\ge t^{15/16}/h+P^+(h)$. Thus, by (7.14) and Lemma 7.2,

(7.17)
$$\begin{aligned}T_2&\le\frac{4r}{\theta}\max_{\substack{1\le\ell\le s\\ \nu(\ell)\le\nu(s)}}\sum_{\substack{P^+(g)\le e^\theta\\ \tau(g)\ge\ell+1}}\frac{\log(\frac{d_{\ell+1}(g)}{d_\ell(g)})}{g}\sum_{h\in\mathscr{P}(e^\theta,te^\eta)}\frac{L(h;\eta)}{h\log^2(t^{15/16}/h+P^+(h))}\\ &\ll_r(\log\eta)^{\nu(s)+1}\sum_{h\in\mathscr{P}(e^\theta,te^\eta)}\frac{L(h;\eta)}{h\log^2(t^{15/16}/h+P^+(h))}.\end{aligned}$$

If $g > t^{1/16}$, we use the bound $t/(gh) + P^+(gh) \ge P^+(g)$. By Lemma 3.1, $L(h;\eta) \le \eta\tau(h)$. Since

$$\sum_{h\in\mathscr{P}(e^\theta,te^\eta)} \frac{\tau(h)}{h} \ll \frac{\log^2(te^\eta)}{\theta^2} \ll_{r,C} \left(\frac{\log t}{\theta}\right)^2,$$

we have, by (7.14),

$$T_3 \ll_{r,C} \left(\frac{\log t}{\theta}\right)^2 \max_{\substack{1\le \ell\le s\\ \nu(\ell)\le\nu(s)}} \sum_{\substack{P^+(g)\le e^\theta\\ \tau(g)\ge \ell+1\\ g>t^{1/16}}} \frac{\log(d_{\ell+1}(g)/d_\ell(g))}{g\log^2 P^+(g)}. \tag{7.18}$$

By Lemma 7.2 and partial summation (with respect to $P^+(g)$), the sum on g in (7.18) is

$$\begin{aligned}
&\ll (\log\theta)^{\nu(\ell)+1}\left[\frac{1}{\theta}e^{-\frac{\log t}{64\theta}} + \int_2^{e^\theta} \frac{e^{-\frac{\log t}{64\log u}}}{u\log^2 u}\,du\right]\\
&\le (\log\theta)^{\nu(\ell)+1}\left(\frac{1}{\theta}+\frac{64}{\log t}\right)e^{-\frac{\log t}{64\theta}}.
\end{aligned}$$

Since $\theta \ll_{r,C} \log t$, (7.18) and (3.6) give

$$\begin{aligned}
T_3 &\ll_{r,C} (\log\eta)^{\nu(s)+1}\frac{\theta}{\log^2 t}\left[\left(\frac{\log t}{\theta}\right)^4 e^{-\frac{\log t}{64\theta}}\right]\\
&\ll_{r,C} (\log\eta)^{\nu(r)+1}\frac{\theta}{\log^2 t} \ll_{r,C} (\log\eta)^{\nu(r)+1}S^*(t;\eta).
\end{aligned}$$

Combining this with (7.13) and (7.17), we obtain

$$S_s(t,\eta) \ll_{r,C} (\log\eta)^{\nu(r)+1}\left(S^*(t,\eta) + \sum_{h\in\mathscr{P}(e^\theta,te^\eta)} \frac{L(h;\eta)}{h\log^2(t^{15/16}/h+P^+(h))}\right).$$

As in the proof of (7.1), write each h in the above sum as $h = a_1a_2$, where $a_1 \in \mathscr{P}(e^\theta, e^\eta)$ and $a_2 \in \mathscr{P}(e^\eta, te^\eta)$. Let T_4 be the contribution to the sum from those h with $a_1 \le t^{1/16}$, and let T_5 be the remaining portion of the sum. By Lemma 3.1, $L(h;\eta) \le \tau(a_1)L(a_2;\eta)$, and we also have

$$\sum_{a_1\in\mathscr{P}(e^\theta,e^\eta)} \frac{\tau(a_1)}{a_1} \ll \left(\frac{\eta}{\theta}\right)^2 \ll_{r,C} 1.$$

Thus $T_4 \ll_{r,C} \hat{S}(t;\eta)$. When $a_1 > t^{1/16}$, Lemma 3.1 gives

$$L(h;\eta) \le \left(1+\frac{\log a_1}{\eta}\right)L(a_2;\eta) \ll_{r,C} \frac{\log a_1}{\eta}L(a_2;\eta).$$

Also, $t^{15/16}/h + P^+(h) \ge P^+(a_1) \ge e^\theta$, so

$$T_5 \ll_{r,C} \frac{1}{\eta^3} \sum_{\substack{a_1 \in \mathscr{P}(e^\theta, e^\eta) \\ a_1 > t^{1/16}}} \frac{\log a_1}{a_1} \sum_{a_2 \in \mathscr{P}(e^\eta, te^\eta)} \frac{L(a_2;\eta)}{a_2}.$$

By Lemma 2.5, the sum on a_1 is

$$\ll_{r,C} \frac{\eta^3}{\log^2 t}\left[\left(\frac{\log t}{\eta}\right)^3 e^{-\frac{\log t}{64\eta}}\right] \ll_{r,C} \frac{\eta^3}{\log^2 t}.$$

Thus, $T_5 \ll_{r,C} \hat{S}(t;\eta)$ and hence

$$S_s(t;\eta) \ll_{r,C} (\log \eta)^{\nu(r)+1}\big(S^*(t;\eta) + \hat{S}(t;\eta)\big).$$

The lemma now follows from (7.3). $\square$

8. Upper bounds: reduction to an integral

In this section we prove Lemma 3.5. We have $T_0(\sigma, P, Q) = 0$ if $Q > 1$ and $T_0(\sigma, P, 1) = \sigma$. Now let $k \ge 1$ and put $\gamma = \frac{1}{\log P}$. We have

$$T_k(\sigma, P, Q) \le Q^{-\gamma} \sum_{\substack{a \in \mathscr{P}^*(e^\sigma, P) \\ \omega(a) = k}} \frac{L(a;\sigma)}{a^{1-\gamma}}.$$

The parameter γ has been chosen so that the sum is only a constant multiplicative factor larger than the corresponding sum with $\gamma = 0$. Since $p^\gamma \le 1 + 2\gamma \log p$ for $p \le P$, we have

$$\sum_{e^\sigma < p \le P} \frac{1}{p^{1-\gamma}} \le \sum_{e^\sigma < p \le P} \frac{1}{p} + 2\gamma \sum_{p \le P} \frac{\log p}{p} \le \log\log P - \log(1+\sigma) + O(1).$$

By an argument similar to that used to construct the sets D_j in Section 4, we find that there is an absolute constant K so that the following holds for all σ, P: the interval $(e^\sigma, P]$ may be partitioned into subintervals $E_0, \dots, E_{v+K-1}$ with v as given in Lemma 3.5 and for each j,

$$\sum_{p \in E_j} \frac{1}{p^{1-\gamma}} \le \log 2$$

and

$$p \in E_j \implies \frac{\log\log p - \log(1+\sigma)}{\log 2} \le j + K. \tag{8.1}$$

Consider $a = p_1 \cdots p_k$, $e^\sigma < p_1 < \cdots < p_k \le P$ and define j_i by $p_i \in E_{j_i}$ $(1 \le i \le k)$. Put $l_i = \frac{\log\log p_i}{\log 2}$. By Lemma 3.1 (v) and (8.1),

$$\begin{aligned} L(a;\sigma) &\le 2^k \min_{0 \le g \le k} 2^{-g}(2^{l_1} + \cdots + 2^{l_g} + \sigma) \\ &\le (\sigma+1)2^{k+K} F(\mathbf{j}), \end{aligned}$$

where

$$F(\mathbf{j}) = \min_{0\le g\le k} 2^{-g}(2^{j_1} + \cdots + 2^{j_g} + \min(1,\sigma)).$$

Let J denote the set of vectors $\mathbf{j}$ satisfying $0 \le j_1 \le \cdots \le j_k \le v + K - 1$. Then

$$T_k(\sigma, P, Q) \le Q^{-\gamma} 2^{k+K} \sum_{\mathbf{j}\in J} F(\mathbf{j}) \sum_{\substack{p_1<\cdots<p_k \\ p_i\in E_{j_i}\ (1\le i\le k)}} \frac{1}{(p_1\cdots p_k)^{1-\gamma}}.$$

If b_j is the number of primes p_i in E_j for $0 \le j \le v + K - 1$, the sum over $p_1, \cdots, p_k$ above is at most

$$\begin{aligned}\prod_{j=0}^{v+K-1} \frac{1}{b_j!}\Bigg(\sum_{p\in E_j} \frac{1}{p^{1-\gamma}}\Bigg)^{b_j} &\le \frac{(\log 2)^k}{b_0!\cdots b_{v+K-1}!} \\ &= ((v+K)\log 2)^k \int_{R(\mathbf{j})} 1\, d\boldsymbol{\xi} \\ &\le e^{10K}(v\log 2)^k \int_{R(\mathbf{j})} 1\, d\boldsymbol{\xi},\end{aligned}$$

where

$$R(\mathbf{j}) = \{0 \le \xi_1 \le \cdots \le \xi_k \le 1 : j_i \le (v+K)\xi_i \le j_i + 1\ \ \forall i\} \subseteq R_k.$$

Finally, since $2^{j_i} \le 2^{(v+K)\xi_i} \le 2^K 2^{v\xi_i}$ for each i,

$$\sum_{\mathbf{j}\in J} F(\mathbf{j}) \int_{R(\mathbf{j})} 1 d\boldsymbol{\xi} \le 2^K U_k(v;\alpha).$$

9. Lower bounds: isolated divisors

In this section we prove Lemmas 4.3 and 4.4.

Proof of Lemma 4.3. Write $a = a' p_1 \cdots p_r$, where

$$a' \le y^{\frac{c'}{100r}}, \qquad y^{\frac{c'}{2r}} < p_1 < \cdots < p_r \le y^{\frac{c'}{r}}.$$

We have

$$\sum_{a\le y^{2c'}} \frac{L_r(a;\eta)}{a} \ge \sum_{a'} \frac{1}{a'} \int \sum_{\substack{p_1,\ldots,p_r \\ \tau(a'p_1\cdots p_r, e^u, e^{u+\eta})=r}} \frac{1}{p_1\cdots p_r}\, du. \tag{9.1}$$

Fix a' and fix u so that $y^{\frac{0.6c'}{r}} \le e^u \le y^{\frac{0.7c'}{r}}$. The measure of such u is $\gg_{r,c'} \log y$. Since $e^{u+\eta} \le y^{\frac{0.8c'}{r}}$, a divisor of $a' p_1 \cdots p_r$ lying in $(e^u, e^{u+\eta})$ must have the form dp_i, where $d|a'$ and $1 \le i \le r$. If

$$\log\left(\frac{e^u}{p_i}\right) \in \mathscr{L}_1(a';\eta) \qquad (1 \le i \le r), \tag{9.2}$$

there are exactly r such divisors. Assume that $I(a';\eta)\ge 1$. If we take, for each i, $e^u/d_i < p_i \le e^{u+\eta}/d_i$ for some η-isolated divisor d_i of a', then (9.2) will be satisfied. By Lemma 2.2, the sum over $p_1,\dots,p_r$ in (9.1) is

$$\ge \frac{1}{r!}\prod_{i=1}^{r}\Bigg(\sum_{\log(e^u/p_i)\in\mathscr{L}_1(a';\eta)}\frac{1}{p_i}\Bigg)-O_{r,c'}\left(\frac{1}{y^{c'/2r}}\right)\gg_{r,c'}\left(\frac{\eta I(a';\eta)}{\log y}\right)^r.$$

By inserting this into (9.1) and using Lemma 4.1, we complete the proof. $\square$

Proof of Lemma 4.4. (i) Write $a=gh$, where $g\le K:=\min(e^\eta,y^{c'})$, $e^{2\eta}<P^-(h)$, $h\le y^{c'}$ and $\tau(g)\ge r+1$. In particular, if $\eta\ge\frac{c'}{2}\log y$, then $h=1$. Let f be a 2η-isolated divisor of h. As before, let $d_j(g)$ be the jth smallest divisor of g. If $fd_r(g)\le e^{u+\eta}\le fd_{r+1}(g)$, then $\tau(gh,e^u,e^{u+\eta})=r$ and hence

$$L_r(gh;\eta)\ge\log\left(\frac{d_{r+1}(g)}{d_r(g)}\right)I(h;2\eta).$$

By Lemma 4.1,

$$H_r(x,y,z)\gg_{r,c',C}\frac{x}{\log^2 y}A_r(K)\sum_{\substack{h\le y^{c'}\\ P^-(h)>e^{2\eta}}}\frac{I(h;2\eta)}{h}, \tag{9.3}$$

where

$$A_r(K)=\sum_{\substack{g\le K\\ \tau(g)\ge r+1}}\frac{\log(d_{r+1}(g)/d_r(g))}{g}.$$

To complete the proof of part (i) of the lemma, it suffices to show

$$A_r(K)\gg_r(\log K)(\log\log K)^{\nu(r)+1}\qquad (K\ge e^{1000\cdot 3^{2r}}) \tag{9.4}$$

and note that $K\ge e^{\eta c'/C}$. We restrict attention to those $g=p_1\cdots p_r m$, with

$$p_1\le K^{1/3^{2r}},\quad p_{j-1}^3<p_j<K^{1/3^{2r-2j+2}}\quad (2\le j\le r), \tag{9.5}$$

$P^-(m)>p_r$ and $m\le K^{1/2}$. Inequality (9.5) implies that

$$\frac{d_{r+1}(g)}{d_r(g)}=\frac{p_v}{p_1\cdots p_{v-1}}\ge p_v^{1/2},\quad v=\nu(r)+1.$$

Hence

$$A\ge\frac12\sum_{\substack{p_1,\dots,p_r\\ (9.5)}}\frac{\log p_v}{p_1\cdots p_r}\sum_{\substack{P^-(m)>p_r\\ m\le K^{1/2}}}\frac{1}{m}.$$

Since $p_r\le K^{1/9}$, by Lemma 2.3 and partial summation, the sum on m is $\gg\frac{\log K}{\log p_r}$. If $v=r$, then $r\in\{1,2\}$ and we obtain

$$A\gg\log K\sum_{p_1,\cdots,p_r}\frac{1}{p_1\cdots p_r}\gg\log K(\log\log K)^r,$$

proving (9.4) in this case. Otherwise, $v < r$. By Lemma 2.2 and Bertrand's Postulate, if $u > w^2 > 4$, then

$$\sum_{w<p\le u} \frac{1}{p\log p} \ge \frac{1}{2\log w} \sum_{w<p\le w^2} \frac{1}{p} \gg \frac{1}{\log w}.$$

Applying this iteratively and using (9.5), we obtain

$$\sum_{p_{v+1},\dots,p_r} \frac{1}{p_{v+1}\cdots p_r \log p_r} \gg_r \frac{1}{\log p_v}.$$

Finally, we have

$$\sum_{p_1,\cdots,p_v} \frac{1}{p_1\cdots p_v} \gg_r (\log\log K)^v = (\log\log K)^{\nu(r)+1}.$$

This is trivially true when $K \le K_0(r)$, for a large constant $K_0(r)$, and for larger K the sum includes $p_1,\dots,p_v$ satisfying

$$p_i \in \Big(\exp[(\log K)^{i/2r}], \exp[(\log K)^{(i+1/2)/2r}]\Big)\ (1\le i\le v)$$

by (9.5). This proves (9.4) in the second case.

(ii) Suppose $y_0(r,c) \le y$ and $y^2 \le z \le x^{1-c}/y$. Consider $n = gpq$, where $g \le y^{c/2}$, $\tau(g) \ge r+1$, p is prime, $yd_r(g)/g < p \le yd_{r+1}(g)/g$, $P^-(q) > z$ and $\frac{x}{2pg} < q \le \frac{x}{pg}$. We have $y \ge p > y/g > g$, so each n has at most one factorization of this type. If $d|n$ and $y < d \le z$, then $p|d$. Thus $\tau(n,y,z) = r$, because

$$\frac{pg}{d_{r+1}(g)} \le y < \frac{pg}{d_r(g)} < \frac{pg}{d_{r-1}(g)} < \cdots < pg \le z.$$

Since $pg \le yx^{c/2} \le \frac{x}{4z}$, Lemma 2.3 implies that for each pair g,p, the number of q is $\gg \frac{x}{pg\log z}$. By Lemma 2.2, for each g,

$$\sum \frac{1}{p} \gg \frac{\log(d_{r+1}(g)/d_r(g))}{\log y}.$$

Using (9.4), the number of such n is

$$\gg \frac{x}{(\log y)(\log z)} A_r(y^{c/2}) \gg_{r,c} \frac{x(\log\log y)^{\nu(r)+1}}{\log z}. \qquad \square$$

10. Lower bounds: reduction to a volume

Proof of Lemma 4.7. Recall the definitions of the sets D_j and numbers λ_j from Section 4. For $j \ge 0$ let $b'_j = \sum_{i\le j} b_j$. Let $a = p_1\cdots p_k$, where $k = b_1 + \cdots + b_h$,

(10.1) $$p_{b'_{j-1}+1},\dots,p_{b'_j} \in D_j \qquad (m \le j \le h)$$

and the primes in each interval D_j are unordered. Observe that $W(p_1 \cdots p_k; \sigma)$ is the number of pairs $Y, Z \subseteq \{1, \dots, k\}$ with

$$\left| \sum_{i \in Y} \log p_i - \sum_{i \in Z} \log p_i \right| \le \sigma. \tag{10.2}$$

We thus have

$$\sum_{a \in \mathscr{A}(\mathbf{b})} \frac{W(a;\sigma)}{a} \le \frac{1}{b_m! \cdots b_h!} \sum_{Y,Z \subseteq \{1,\dots,k\}} \sum_{\substack{p_1,\dots,p_k \\ (10.1),(10.2)}} \frac{1}{p_1 \cdots p_k}. \tag{10.3}$$

The contribution from terms with $Y = Z$ is

$$\frac{2^k}{b_m! \cdots b_h!} \sum_{\substack{p_1,\dots,p_k \\ (10.1)}} \frac{1}{p_1 \cdots p_k} \le \frac{(2\log 2)^k}{b_m! \cdots b_h!}. \tag{10.4}$$

If $Y \ne Z$, let $I = \max[(Y \cup Z) - (Y \cap Z)]$ and define $E(I)$ by $p_I \in D_{E(I)}$, i.e., $b'_{E(I)-1} < I \le b'_{E(I)}$. Let

$$\ell = \min\{j : \lambda_j \ge \sigma^{-2}\}.$$

We distinguish two cases: (i) $E(I) > \ell$; (ii) $m \le E(I) \le \ell$.

In case (i), fix all of the p_i except for p_I. Inequality (10.2) implies that $U \le p_I \le e^{2\sigma} U$ for some number $U \ge \lambda_{E(I)}$. If $\sigma \ge 1$, then by (2.2),

$$\sum_{U \le p_I \le e^{2\sigma}U} \frac{1}{p_I} \ll \log\left(1 + \frac{2\sigma}{\log U}\right) \ll \frac{\sigma}{\log U} \ll \sigma 2^{-E(I)}.$$

If $\sigma < 1$, Lemma 2.1 implies

$$\sum_{U \le p_I \le e^{2\sigma}U} \frac{1}{p_I} \ll \frac{\sigma}{\log(\sigma U)} \ll \frac{\sigma}{\log U} \ll \sigma 2^{-E(I)},$$

where the second inequality follows from $U \ge \lambda_\ell \ge \sigma^{-2}$. Thus, by (4.3) the inner sum on the right side of (10.3) is $\ll \sigma 2^{-E(I)} (\log 2)^k$. With I fixed, there correspond $2^{k-I+1} 4^{I-1} = 2^{k+I-1}$ pairs Y, Z. We find that the contribution to the right side of (10.3) from those Y, Z counted in case (i) is

$$\begin{aligned} &\ll \frac{\sigma(2\log 2)^k}{b_m! \cdots b_h!} \sum_{I=1}^{k} 2^{I-E(I)} \\ &\ll \frac{\sigma(2\log 2)^k}{b_m! \cdots b_h!} \sum_{j=m}^{h} 2^{-j} \sum_{b'_{j-1} < I \le b'_j} 2^I \ll \frac{\sigma(2\log 2)^k}{b_m! \cdots b_h!} \sum_{j=m}^{h} 2^{-j+b_m+\cdots+b_j}. \end{aligned} \tag{10.5}$$

For case (ii), note that $\sigma < 1$ and write

$$a = a' p_{b'_\ell+1} \cdots p_k, \qquad a' = p_1 \cdots p_{b'_\ell}.$$

By hypothesis, $Y \cap \{b'_\ell + 1, \dots, k\} = Z \cap \{b'_\ell + 1, \dots, k\}$. By (4.3), the contribution to the right side of (10.3) from those Y, Z counted in case (ii) is

$$\le \frac{(2\log 2)^{b_{\ell+1}+\cdots+b_h}}{b_{\ell+1}! \dots b_h!} \sum_{a'} \frac{W(a';\sigma) - \tau(a')}{a'}.$$

Suppose $d_1|a'$, $d_2|a'$ and $1 < d_2/d_1 \le e^\sigma$. Let $d = (d_1, d_2)$, $d_1 = f_1 d$, $d_2 = f_2 d$ and $a' = df_1f_2a''$. Then

$$\begin{aligned}
\sum_{a'} \frac{W(a';\sigma) - \tau(a')}{a'} &\le 2 \sum_{P^+(a''df_1) \le \lambda_\ell} \frac{1}{a''df_1} \sum_{f_1 < f_2 \le e^\sigma f_1} \frac{1}{f_2} \\
&\le 4\sigma \sum_{P^+(a''df_1) \le \lambda_\ell} \frac{1}{a''df_1} \\
&= 4\sigma \prod_{p \le \lambda_\ell} \left(1 + \frac{1}{p}\right)^3 \\
&\le 4\sigma \exp\left(3 \sum_{p \le \lambda_\ell} \frac{1}{p}\right) \le 2^{3\ell+2}\sigma.
\end{aligned}$$

Since $\sigma > \lambda_{\ell-1}^{-1/2}$, we have $2^{3\ell+2}\sigma < \exp\{-2^{2\ell/3}\}$ if m is large. On the other hand, we have assumed that $b_j \le 2^{j/2}$ for every j, and thus

$$\frac{1}{b_m! \cdots b_\ell!} \ge \left(2^{(\ell/2)\cdot 2^{\ell/2}}\right)^{-\ell} \ge 100 \exp\{-2^{2\ell/3}\} > 100 \cdot 2^{3\ell+2}\sigma$$

for large m. Therefore, the contribution to the right side of (10.3) from those Y, Z counted in case (ii) is

$$\le 0.01 \frac{(2\log 2)^k}{b_m! \cdots b_h!}. \tag{10.6}$$

Together, (10.3), (10.4), (10.5), and (10.6) imply that

$$\sum_{a \in \mathscr{A}(\mathbf{b})} \frac{W(a;\sigma)}{a} \le \frac{(2\log 2)^k}{b_m! \dots b_h!} \left[1.01 + O\left(\sigma \sum_{j=m}^{h} 2^{-j+b_m+\cdots+b_j}\right)\right]. \qquad \square$$

Proof of Lemma 4.8. Let $m = M + \max\left(0, \left\lfloor \frac{\log\sigma}{\log 2} \right\rfloor\right)$, put $h = v + m - 1$ and suppose $\mathbf{b}$ satisfies $b_j = 0$ $(1 \le j \le m-1)$, $b_1 + \cdots + b_h = k$ and

$$\text{for all } j \ge 1,\ b_j \le 2^{j/10} \text{ and } b_{h+1-j} \le 2^{(M+j)/10}. \tag{10.7}$$

Since $h \le \frac{\log\log y}{\log 2} - M - 1/\alpha$,

$$
\begin{aligned}
\sum_j b_j 2^j &\le 2^{h+\frac{M+1}{10}}(1+2^{-0.9}+2^{-1.8}+\cdots)\\
&\le 2^{h+M/10+2}\\
&\le \frac{\log y}{2^{\frac{9}{10}M-2+1/\alpha}} \le \frac{\alpha \log y}{2^{c_3+c_4}}.
\end{aligned}
$$

Using Lemma 4.6, it follows that $a \le y^\alpha$ for $a \in \mathscr{A}(\mathbf{b})$. Also $P^-(a) > \lambda_{m-1} > e^\sigma$. By the definition of the sets D_j,

$$
\begin{aligned}
\sum_{a\in\mathscr{A}(\mathbf{b})} \frac{1}{a} &= \prod_{j=m}^h \frac{1}{b_j!}\Bigg(\sum_{p_1\in D_j}\frac{1}{p_1}\sum_{\substack{p_2\in D_j\\ p_2\ne p_1}}\frac{1}{p_2}\cdots \sum_{\substack{p_{b_j}\in D_j\\ p_{b_j}\notin\{p_1,\dots,p_{b_j-1}\}}}\frac{1}{p_{b_j}}\Bigg)\\
&\ge \prod_{j=m}^h \frac{1}{b_j!}\Bigg(\sum_{p\in D_j}\frac{1}{p} - \frac{b_j-1}{\lambda_{j-1}}\Bigg)^{b_j}\\
&\ge \prod_{j=m}^h \frac{1}{b_j!}\Bigg(\log 2 - \frac{b_j}{\lambda_{j-1}}\Bigg)^{b_j} \qquad (10.8)\\
&\ge \frac{(\log 2)^k}{b_m!\cdots b_h!}\prod_{j=m}^h\left(1-\frac{2^{j/10}}{\exp\{2^{j-1+c_3-c_4}\}}\right)^{2^{j/10}}\\
&\ge 0.999\frac{(\log 2)^k}{b_m!\cdots b_h!}.
\end{aligned}
$$

By Lemma 4.7 and (10.8),

$$
\sum_{a\in\mathscr{A}(\mathbf{b})}\frac{3\tau(a)-2W(a;\sigma)}{a} \ge \frac{(2\log 2)^k}{b_m!\cdots b_h!}\left[0.9 - 2^{c_5+1}\sigma\sum_{j=m}^h 2^{-j+b_m+\cdots+b_j}\right].
$$

For $1\le i\le v$, set $g_i = b_{m-1+i}$. We have $2^{(M+i-1)/10} \ge M+i^2$ for $M\ge 200$ and $i\ge 1$, so (10.7) is implied by

$$
\forall i\ge 1, g_i \le M+i^2 \text{ and } g_{v+1-i}\le M+i^2. \qquad (10.9)
$$

Assume in addition that

$$
2^{m-1}\sum_{j=m}^h 2^{-j+b_m+\cdots+b_j} = \sum_{i=1}^v 2^{-i+g_1+\cdots+g_i} \le 2^{s+1}. \qquad (10.10)
$$

Since $2^{s-m+2} = \frac{1}{\sigma}2^{-c_5-8}$, we obtain

$$
\sum_{a\in\mathscr{A}(\mathbf{b})}\frac{3\tau(a)-2W(a;\sigma)}{a} \ge \frac{(2\log 2)^k}{3g_1!\cdots g_v!}. \qquad (10.11)
$$

Let $\mathscr{G}$ be the set of $\mathbf{g} = (g_1, \cdots, g_v)$ with $g_1 + \cdots + g_v = k$ and also satisfying (10.9) and (10.10). We have

$$\frac{1}{g_1! \cdots g_v!} = \mathrm{Vol}(R(\mathbf{g})), \tag{10.12}$$

where $R(\mathbf{g})$ is the set of $\mathbf{x} \in \mathbb{R}^k$ with $0 \le x_1 \le \cdots \le x_k < v$ and exactly g_i of the variables x_j lie in $[i-1, i)$ for each i. We claim that

$$\mathrm{Vol}(\cup_{\mathbf{g}\in\mathscr{G}} R(\mathbf{g})) \ge v^k \,\mathrm{Vol}(Y_k(s, v)). \tag{10.13}$$

Take $\boldsymbol{\xi} \in Y_k(s, v)$ with $\xi_k < 1$ and let $x_j = v\xi_j$ for each j. Let g_i be the number of x_j lying in $[i-1, i)$. By condition (ii) in the definition of $Y_k(s, v)$,

$$x_{M+i^2} > i \quad \text{and} \quad x_{k+1-(M+i^2)} < v - i \qquad (1 \le i \le \sqrt{k-M}),$$

which implies (10.9). Condition (iii) in the definition of $Y_k(s, v)$ implies

$$2^s \ge \sum_{j=1}^{k} 2^{j-x_j} \ge \sum_{i=1}^{v} 2^{-i} \sum_{j: x_j \in [i-1, i)} 2^j \ge \sum_{i=1}^{v} 2^{-i+g_1+\cdots+g_i-1}.$$

Because (10.9) and (10.10) hold, $\mathbf{g} \in \mathscr{G}$ and $\mathbf{x} \in R(\mathbf{g})$. This proves the claim (10.13). Finally, combining (10.11), (10.12) and (10.13) proves the lemma. □

11. Uniform order statistics

Let $X_1, \ldots, X_k$ be independent, uniformly distributed random variables in $[0, 1]$, and let $\xi_1, \ldots, \xi_k$ be their order statistics, so that $0 \le \xi_1 \le \cdots \le \xi_k \le 1$. Our main interest in this section is to estimate $Q_k(u, v)$, the probability that $\xi_i \ge \frac{i-u}{v}$ for every i. The special case $Q_k(\lambda\sqrt{k}, k)$ is the distribution function (as a function of λ) of the quantity

$$D_k^+ = -\sqrt{k} \inf_{1 \le i \le k} (\xi_i - i/k) = \sqrt{k} \sup_{0 \le y \le 1} \left(k^{-1} \sum_{X_i \le y} 1 - y \right),$$

which is one of the Kolmogorov-Smirnov statistics. In 1939, N. V. Smirnov proved [36] for each fixed $\lambda \ge 0$ the asymptotic formula

$$Q_k(\lambda\sqrt{k}, k) \sim 1 - e^{-2\lambda^2} \qquad (k \to \infty).$$

We need bounds on $Q_k(u, v)$ which are uniform in k, u and v, particularly when u is small or $u + v - n$ is small.

Lemma 11.1. *For a given pair u, v, let $w = w(u, v) = u+v-n$. Uniformly in $u > 0$, $w > 0$ and $k \ge 1$,*

$$Q_k(u, v) \ll \frac{(u+1)(w+1)^2}{k}. \tag{11.1}$$

If $k \ge 1$ *and* $1 \le u \le k$*, then*

$$Q_k(u, k+1-u) \ge \frac{u-1/2}{k+1/2}. \tag{11.2}$$

Remarks. The more precise estimate

$$Q_k(u,v) = 1 - e^{-2uw/k} + O\left(\frac{u+w}{k}\right) \qquad (k \ge 1, u \ge 0, w \ge 0),$$

which was proved in an earlier version of this paper, has a much longer proof which will be given in a separate paper [15].

Proof. If $\min_{1\le i\le k}(\xi_i - \frac{i-u}{v}) \le 0$, let l be the smallest index with $\xi_l \le \frac{l-u}{v}$ and write $\xi_l = \frac{l-u-\lambda}{v}$, so that $0 \le \lambda \le 1$. Also let

$$R_l(\lambda) = \operatorname{Vol}\left\{0 \le \xi_1 \le \cdots \le \xi_{l-1} \le \frac{l-u-\lambda}{v} : \xi_i \ge \frac{i-u}{v}\,(1 \le i \le l-1)\right\}.$$

Then

$$\begin{aligned} Q_k(u,v) &= 1 - \frac{k!}{v}\int_0^1 \sum_{u+\lambda\le l\le k} R_l(\lambda)\operatorname{Vol}\left\{\frac{l-u-\lambda}{v} \le \xi_{l+1} \le \cdots \le \xi_k \le 1\right\} d\lambda \\ &= 1 - \frac{k!}{v}\int_0^1 \sum_{u+\lambda\le l\le k} \frac{R_l(\lambda)}{(k-l)!}\left(\frac{k+w+\lambda-l}{v}\right)^{k-l} d\lambda. \end{aligned}$$

Let $0 \le a \le k-u$ and suppose that $\xi_k \le 1 - \frac{w+a}{v} = \frac{k-u-a}{v}$. Then $\min_{1\le i\le k}\xi_i - \frac{i-u}{v} \le 0$. Defining l and λ as before, we have

$$\begin{aligned} \left(1 - \frac{w+a}{v}\right)^k &= k!\operatorname{Vol}\left\{0 \le \xi_1 \le \cdots \le \xi_k \le 1 - \frac{w+a}{v}\right\} \\ &= \frac{k!}{v}\int_0^1 \sum_{u+\lambda\le l\le k-a+\lambda} \frac{R_l(\lambda)}{(k-l)!}\left(\frac{k-l-a+\lambda}{v}\right)^{k-l} d\lambda. \end{aligned}$$

Thus, for any $A > 0$,

$$\begin{aligned} Q_k(u,v) = 1 - A\left(1 - \frac{w+a}{v}\right)^k \\ - \frac{k!}{v}\int_0^1 \sum_{k-a+\lambda<l\le k} \frac{R_l(\lambda)}{(k-l)!}\left(\frac{k+w+\lambda-l}{v}\right)^{k-l} d\lambda \\ + \frac{k!}{v}\int_0^1 \sum_{l=\lfloor u+\lambda+1\rfloor}^{\lfloor k-a+\lambda\rfloor} \frac{R_l(\lambda)}{(k-l)!v^{k-l}}\left[A(k-l-a+\lambda)^{k-l} - (k-l+w+\lambda)^{k-l}\right] d\lambda. \end{aligned} \tag{11.3}$$

To prove (11.1), we may assume without loss of generality that $k \ge 10$, $u \le k/10$ and $w \le \sqrt{k}$. Let $a = 2w+2$, and note that $2-\lambda \ge \lambda$. Now,

$$\begin{aligned}\left(\frac{k-l-w-2+\lambda}{k-l+w+\lambda}\right)^{k-l} &= \left(1-\frac{w+2-\lambda}{k-l}\right)^{k-l}\left(1+\frac{w+\lambda}{k-l}\right)^{-(k-l)} \\ &= \exp\left\{-(2w+2)+\sum_{j=2}^{\infty}\frac{-(w+2-\lambda)^j+(-1)^j(w+\lambda)^j}{j(k-l)^{j-1}}\right\} \\ &\le e^{-(2w+2)}.\end{aligned}$$

Thus, taking $A=e^{2w+2}$ in (11.3), we conclude that

$$\begin{aligned}Q_k(u,v) &\le 1-e^{2w+2}\left(1-\frac{2w+2}{v}\right)^k \\ &= 1-\exp\left\{\frac{2w+2}{v}\left(v-k+O(w)\right)\right\} \\ &= 1-\exp\left\{\frac{-2uw+O(u+w^2+1)}{v}\right\} \\ &\le \frac{2uw+O(u+w^2+1)}{v} \ll \frac{(u+1)(w+1)^2}{k}.\end{aligned}$$

To prove (11.2), let $a=0$ in (11.3). The first sum on l in (11.3) is empty, and

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

Thus, taking $A=e^w$ in (11.3), we obtain

$$Q_k(u,v) \ge 1-e^w\left(1-\frac{w}{v}\right)^k.$$

Now let $v=k+1-u$ (so that $w=1$) and let $f(x)=e^x(1-\frac{x}{k+1-u})^k$. For $0\le x\le 1$,

$$f'(x) = -f(x)\frac{u-1+x}{k+1-u-x} \le 0;$$

hence

$$\begin{aligned}f(1) &= 1-\int_0^1 f(x)\ \frac{u-1+x}{k+1-u-x}\,dx \\ &\le 1-\frac{f(1)}{k+1-u}\int_0^1 u-1+x\,dx = 1-\frac{u-1/2}{k+1-u}f(1).\end{aligned}$$

This implies that $f(1)\le\frac{k+1-u}{k+1/2}$ and (11.2) follows. □

12. The lower bound volume

In this section we prove Lemma 4.9 using Lemma 11.1. We begin with a crude upper bound on a combinatorial sum which will also be needed for the upper bound integral in Section 13.

Lemma 12.1. *Let* $0 < \varepsilon \le 1$, $t \ge 10$, *and suppose* a, b *are real numbers with* $a + b > -(1-\varepsilon)t$. *Then*

$$\sum_{\substack{1\le j\le t-1\\ -a<j<b+t}} \binom{t}{j}(a+j)^{j-1}(b+t-j)^{t-j-1} \le e^{5+1/\varepsilon}(t+a+b)^{t-1}.$$

Proof. Let $C_t(a,b)$ denote the sum in the lemma. Since $C_t(a,b) = C_t(b,a)$, we may suppose $a \le b$. We also assume that $a > 1-t$ and $b > 1-t$, otherwise $C_t(a,b) = 0$. The associated "complete" sum is evaluated exactly using one of Abel's identities ([34, p. 20, Eq. (20)])

$$\sum_{j=0}^{t} \binom{t}{j}(a+j)^{j-1}(b+t-j)^{t-j-1} = \left(\frac{1}{a}+\frac{1}{b}\right)(t+a+b)^{t-1} \qquad (ab \ne 0). \tag{12.1}$$

Note that $C_t(a,b)$ is increasing in a and in b. If $-1 \le a \le b$, put $A = \max(1/2, a)$ and $B = \max(1/2, b)$. By (12.1),

$$\begin{aligned} C_t(a,b) \le C_t(A,B) &\le \left(\frac{1}{A}+\frac{1}{B}\right)(t+A+B)^{t-1} \\ &\le 4(t+a+b+3)^{t-1} \\ &\le 4e^{\frac{3(t-1)}{t+a+b}}(t+a+b)^{t-1} < e^5(t+a+b)^{t-1}. \end{aligned} \tag{12.2}$$

Next assume $a < -1 \le b$. Since $(1-1/x)^x$ is an increasing function for $x > 1$, when $j > -a$ we have

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

Also, $(a+1)b \le (1-\varepsilon)t < \frac{1-\varepsilon}{\varepsilon}(t+a+b)$. Thus, by (12.2),

$$\begin{aligned} C_t(a,b) &\le \left(\frac{t+a}{t-1}\right)^{t-1} C_t(-1,b) \\ &\le e^5\left(\frac{(t+a)(t+b-1)}{t-1}\right)^{t-1} = e^5\left(t+a+b+\frac{(a+1)b}{t-1}\right)^{t-1} \\ &< e^{4+1/\varepsilon}(t+a+b)^{t-1}. \end{aligned}$$

Lastly, suppose $a \le b < -1$. By (12.2),

$$\begin{aligned} C_t(a,b) &\le \max_{-a<j<b+t}\left(\frac{j+a}{j-1}\right)^{j-1}\left(\frac{b+t-j}{t-j-1}\right)^{t-j-1} C_t(-1,-1) \\ &< e^5(t-2)^{t-1}\max_{-a<j<b+t}\left(\frac{j+a}{j-1}\right)^{j-1}\left(\frac{b+t-j}{t-j-1}\right)^{t-j-1}. \end{aligned}$$

As a function of real j, the maximum above occurs when $\frac{j+a}{j-1} = \frac{b+t-j}{t-j-1}$; that is, $j = \frac{(t-1)(a+1)+b+1}{a+b+2}$. By our assumptions on a, b and t, it is clear that

$1 < j < t-1$. Hence $j+a$ and $b+t-j$ have the same sign, so that $-a < j < b+t$. Then

$$\begin{aligned} C_t(a,b) &\le e^5 \left(\frac{t+a+b}{t-2} \right)^{t-2} (t-2)^{t-1} \\ &< \varepsilon^{-1} e^5 (t+a+b)^{t-1} \\ &< e^{5+1/\varepsilon} (t+a+b)^{t-1}. \end{aligned}$$

□

Proof of Lemma 4.9. For brevity, write

$$S_k(u,v) = \{\boldsymbol{\xi} : 0 \le \xi_1 \le \cdots \le \xi_k \le 1 : \xi_i \ge \tfrac{i-u}{v}\ \forall i\}, \tag{12.3}$$

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

$$F_{k,v}(\boldsymbol{\xi}) = \sum_{j=1}^{k} 2^{j - v\xi_j}. \tag{12.4}$$

Assume v, k, s satisfy the hypotheses of Lemma 4.9, and put $u = k+1-v$. For $1 \le a \le k$, $0 \le b \le k$, let

$$\begin{aligned} V_1(a,b) &= \mathrm{Vol}\{\boldsymbol{\xi} \in S_k(u,v) : \xi_a \le b/v\}, \\ V_2(a,b) &= \mathrm{Vol}\{\boldsymbol{\xi} \in S_k(u,v) : \xi_{k+1-a} \ge 1 - b/v\}. \end{aligned}$$

Make the change of variables $\theta_i = \xi_{a+i}$, so that $\theta_i \ge \frac{i-(u-a)}{v}$ $(1 \le i \le k-a)$. By Lemma 11.1, we have

$$\begin{aligned} V_1(a,b) &\le \mathrm{Vol}\{0 \le \xi_1 \le \cdots \le \xi_a \le \tfrac{b}{v}\}\,\mathrm{Vol}(S_{k-a}(u-a,v)) \\ &= \frac{(b/v)^a}{a!(k-a)!} Q_{k-a}(u-a,v) \\ &\le \frac{b^a}{a!k!} \left(\frac{k}{v} \right)^a Q_{k-a}(u-a,v) \\ &\ll \frac{(100b)^a u}{a!k!(k-a+1)}. \end{aligned} \tag{12.5}$$

Similarly,

$$\begin{aligned} V_2(a,b) &\le \mathrm{Vol}(S_{k-a}(u,v))\,\mathrm{Vol}\{1 - \tfrac{b}{v} \le \xi_{k+1-a} \le \cdots \le \xi_k \le 1\} \\ &\le \frac{(b/v)^a}{a!(k-a)!} Q_{k-a}(u,v) \\ &\ll \frac{a^2 (100b)^a u}{a!k!(k-a+1)}. \end{aligned} \tag{12.6}$$

Next, we show that $F_{k,v}(\boldsymbol{\xi})$ is $\ll 2^u$ on average for $\boldsymbol{\xi} \in S_k(u,v)$. We integrate each term in the sum (12.4) separately, introducing $y = v\xi_j - j + u$, so that

$$\max(0, u-j) \le y \le u + v - j$$

and

$$0 \le \xi_1 \le \cdots \le \xi_{j-1} \le \tfrac{j-u+y}{v} \le \xi_{j+1} \le \cdots \le \xi_k \le 1.$$

Let

$$\begin{aligned} \theta_i &= \tfrac{v\xi_i}{j-u+y} \quad (1 \le i \le j-1), \\ \zeta_i &= \tfrac{v}{u+v-j-y}\left(\xi_{j+i} - \tfrac{j-u+y}{v}\right) \quad (1 \le i \le k-j). \end{aligned}$$

Then $\boldsymbol{\theta} \in S_{j-1}(u, j-u+y)$ and $\boldsymbol{\zeta} \in S_{k-j}(y, u+v-j-y)$. By Lemma 11.1, we obtain

$$\begin{aligned} \int\limits_{S_k(u,v)} F_{k,v}(\boldsymbol{\xi})\, d\boldsymbol{\xi} &= \frac{2^u}{v} \sum_{j=1}^{k} \int_{\max(0,u-j)}^{u+v-j} 2^{-y} \left(\frac{j-u+y}{v}\right)^{j-1} \frac{Q_{j-1}(u, j-u+y)}{(j-1)!} \\ &\qquad \times \left(\frac{u+v-j-y}{v}\right)^{k-j} \frac{Q_{k-j}(y, u+v-j-y)}{(k-j)!}\, dy \\ &\ll \frac{2^u u}{v^k (k+1)!} \int_0^{u+v-1} (y+1)^3 2^{-y} \sum_{\substack{1 \le j \le k \\ u-y<j<u+v-y}} \binom{k+1}{j} \\ &\qquad \times (j+y-u)^{j-1}(u+v-y-j)^{k-j}\, dy. \end{aligned}$$

Since $k \le 100v - 100$, $-u = v-1-k \ge -0.99k$. By Lemma 12.1 (with $t = k+1$, $a = y-u$, $b = -y$ and $\varepsilon = 0.01$), for each y the sum on j on the right side is $\ll v^k$. Since $\int_0^\infty (y+1)^3 2^{-y}\, dy = O(1)$, we conclude that

$$\int\limits_{S_k(u,v)} F_{k,v}(\boldsymbol{\xi})\, d\boldsymbol{\xi} \ll \frac{2^u u}{(k+1)!}. \tag{12.7}$$

Next,

$$\begin{aligned} \mathrm{Vol}(Y_k(s,v)) \ge{}& \mathrm{Vol}\{\boldsymbol{\xi} \in S_k(u,v) : F_{k,v}(\boldsymbol{\xi}) \le 2^s\} \\ &- \sum_{1 \le i \le \sqrt{k-M}} \big(V_1(M+i^2, i) + V_2(M+i^2, i)\big). \end{aligned} \tag{12.8}$$

By (12.5), (12.6) and the simple inequality $h! > (h/e)^h$,

$$\begin{aligned} \sum_{j=1}^{2} \sum_{1 \le i \le \sqrt{k-M}} V_j(M+i^2, i) &\ll \frac{u}{k!} \sum_{1 \le i \le \sqrt{k-M}} \frac{(M+i^2)^2 (100i)^{M+i^2}}{(M+i^2)!(k+1-M-i^2)} \\ &\ll \frac{u}{2^M \cdot (k+1)!}. \end{aligned} \tag{12.9}$$

Also,

$$\mathrm{Vol}\{\boldsymbol{\xi} \in S_k(u,v) : F_{k,v}(\boldsymbol{\xi}) \le 2^s\} \ge \mathrm{Vol}(S_k(u,v)) - 2^{-s} \int\limits_{S_k(u,v)} F_{k,v}(\boldsymbol{\xi})\, d\boldsymbol{\xi}$$

and, by Lemma 11.1,

$$\mathrm{Vol}(S_k(u,v)) \ge \frac{u}{3(k+1)!}.$$

Combining this with (12.7), (12.8) and (12.9), we conclude that

$$\mathrm{Vol}(Y_k(s,v)) \ge \frac{u}{(k+1)!}\left(\frac{1}{4} - \frac{K_1}{2^M} - K_2 2^{u-s}\right),$$

where K_1 and K_2 are positive absolute constants. By hypothesis, $u = k+1-v \le s - M/3$, and the lemma follows provided M is large enough. □

13. The upper bound integral

The purpose of this section is to use the bounds on uniform order statistics proved in Section 11 to prove Lemma 3.6. The primary tools are bounds for the volumes of subsets of $S_k(u,v)$ (defined in (12.3)) in which one or more coordinates is abnormally small or large.

LEMMA 13.1. *Suppose $g,k,t,u,v \in \mathbb{Z}$ satisfy*

$$2 \le g \le k/2,\ t \ge -1,\ v \ge k/10,\ u \ge 0,\ u+v \ge k+1.$$

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

$$\frac{l-u}{v} \le \xi_l \le \frac{l-u+1}{v}, \qquad \xi_{l-g} \ge \frac{l-u-t}{v}. \tag{13.1}$$

Then

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

Proof. Fix l satisfying $\max(u, g+1) \le l \le k$. Let R_l be the subset of $\boldsymbol{\xi} \in S_k(u,v)$ satisfying (13.1) for this particular l. We have $\mathrm{Vol}(R_l) \le V_1V_2V_3V_4$, where by Lemma 11.1,

$$\begin{aligned}
V_1 &= \mathrm{Vol}\{0 \le \xi_1 \le \cdots \le \xi_{l-g-1} \le \tfrac{l-u+1}{v} : \xi_i \ge \tfrac{i-u}{v}\ \forall i\} \\
&= \left(\frac{l-u+1}{v}\right)^{l-g-1} \frac{Q_{l-g-1}(u, l-u+1)}{(l-g-1)!} \\
&\ll \left(\frac{l-u+1}{v}\right)^{l-g-1} \frac{(u+1)g^2}{(l-g)!}, \\
V_2 &= \mathrm{Vol}\{\tfrac{l-u-t}{v} \le \xi_{l-g} \le \cdots \le \xi_{l-1} \le \tfrac{l-u+1}{v}\} = \frac{1}{g!}\left(\frac{t+1}{v}\right)^g, \\
V_3 &= \mathrm{Vol}\{\tfrac{l-u}{v} \le \xi_l \le \tfrac{l-u+1}{v}\} = \frac{1}{v}, \\
V_4 &= \mathrm{Vol}\{\xi_{l+1} \le \cdots \le \xi_k \le 1 : \xi_i \ge \tfrac{i-u}{v}\ \forall i\} \\
&= \frac{1}{(k-l)!}\left(\frac{u+v-l}{v}\right)^{k-l} Q_{k-l}(0, u+v-l)
\end{aligned}$$

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

Thus

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

By Lemma 12.1 (with $t=k-g$, $a=g+1-u$, $b=u+v-k$ and $\varepsilon=\frac{1}{10}$), the sum on l is

$$\ll (v+1)^{k-g-1} \ll \frac{v^{k-g}}{k} \le \frac{v^k(k-g)!}{k\cdot k!}\left(\frac{k}{v}\right)^g \ll \frac{v^k 10^g(k-g)!}{(k+1)!},$$

and the lemma follows. □

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

$$\begin{aligned}\mathscr{T}(k,v,\gamma) = \{\boldsymbol{\xi}\in\mathbb{R}^k : 0\le \xi_1\le\cdots\le\xi_k\le 1,& \\ 2^{v\xi_1}+\cdots+2^{v\xi_j}\ge 2^{j-\gamma}\ (1\le j\le k)\}.&\end{aligned} \tag{13.2}$$

Lemma 13.2. *Suppose k, v, γ are integers with $1 \le k \le 10v$ and $\gamma \ge 0$. Set $b = k-v$. Then*

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

where

$$Y = \begin{cases} b & \text{if } b\ge\gamma+5\\ (\gamma+5-b)^2(\gamma+1) & \text{if } b\le\gamma+4.\end{cases}$$

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

$$\xi_j > \tfrac{j-\gamma-r}{v} \quad (1\le j\le k) \tag{13.3}$$

or

$$\min_{1\le j\le k}(\xi_j - \tfrac{j-\gamma}{v}) = \xi_l - \tfrac{l-\gamma}{v} \in [\tfrac{-h}{v}, \tfrac{1-h}{v}] \quad \text{for some integers } h\ge r+1, 1\le l\le k. \tag{13.4}$$

Let V_1 be the volume of $\boldsymbol{\xi}\in\mathscr{T}(k,v,\gamma)$ satisfying (13.3). If $b\ge\gamma+5$, (13.3) is not possible, so $b\le\gamma+4$ and $r=5$. By Lemma 11.1,

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

If (13.4) holds, then there is an integer m satisfying

$$m\ge h-3,\ 2^m < \tfrac{l}{2},\ \xi_{l-2^m} \ge \tfrac{l-\gamma-2m}{v}. \tag{13.5}$$

To see (13.5), suppose such an m does not exist. Then

$$\begin{aligned}
2^{v\xi_1}+\cdots+2^{v\xi_l} &\le 2\sum_{l/2<j\le l} 2^{v\xi_j} \\
&< 2\Bigg(2^{h-3}2^{l-\gamma-h+1}+\sum_{m\ge h-3} 2^m 2^{l-\gamma-2m}\Bigg) \\
&\le 2^{l-\gamma},
\end{aligned}$$

a contradiction. Let V_2 be the volume of $\boldsymbol{\xi}\in\mathscr{T}(k,v,\gamma)$ satisfying (13.4). Fix h and m satisfying (13.5) and apply Lemma 13.1 with $u=\gamma+h$, $g=2^m$, $t=2m$. The volume of such $\boldsymbol{\xi}$ is

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

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

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

Proof of Lemma 3.6. Assume $k\ge 1$, since the lemma is trivial when $k=0$. Put $b=k-v$ and define

$$F(\boldsymbol{\xi})=\min_{0\le j\le k} 2^{-j}\left(2^{v\xi_1}+\cdots+2^{v\xi_j}+\alpha\right).$$

Let $t=\left\lfloor\frac{\log\alpha}{\log 2}\right\rfloor\le 0$. For integers $m\ge 0$, consider $\boldsymbol{\xi}\in R_k$ satisfying $2^{-m}\alpha\le F(\boldsymbol{\xi})<2^{1-m}\alpha$. For $1\le j\le k$,

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

so that $\boldsymbol{\xi}\in\mathscr{T}(k,v,m+1-t)$. Hence, by Lemma 13.2,

$$\begin{aligned}
U_k(v;\alpha) &\le \sum_{m\ge 0} 2^{1-m}\alpha\,\mathrm{Vol}(\mathscr{T}(k,v,m+1-t)) \\
&\ll \frac{\alpha}{(k+1)!}\sum_{m\ge 0}\frac{2^{-m}Y_m}{2^{2^{b+t-m-1}}}, \\
Y_m &= \begin{cases} b & \text{if } m\le b+t-6 \\ (m+6-t-b)^2(m+2-t) & \text{if } m\ge b+t-5. \end{cases}
\end{aligned}$$

Next,

$$\sum_{m\ge 0}\frac{2^{-m}Y_m}{2^{2^{b+t-m-1}}}=\sum_{0\le m\le b+t-6}\frac{b}{2^m2^{2^{b+t-m-1}}}$$

$$+\sum_{m\ge \max(0,b+t-5)} \frac{(m+6-t-b)^2(m+2-t)}{2^m}.$$

If $b \ge 6-t$, each sum on the right side is $\ll b2^{-b-t}$. If $b \le 5-t$, the first sum is empty and the second is $\ll (6-t-b)^2(2-t)$. In both cases

$$\sum_{m\ge 0} \frac{2^{-m}Y_m}{2^{2^{b+t-m-1}}} \ll \frac{(1+|b+t|^2)(1-t)}{2^{b+t}+1},$$

whence

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

Sometimes this is worse than the simpler bound

$$U_k(v;\alpha) \ll \frac{\alpha}{k!(\alpha 2^{k-v}+1)}, \tag{13.6}$$

which we now prove. When $\alpha 2^{k-v} \le 1$, (13.6) follows from the trivial bound $U_k(v;\alpha) \le \alpha/k!$. Otherwise,

$$\begin{aligned} U_k(v;\alpha) &\le 2^{1-k}\int_{R_k} 2^{v\xi_1}+\cdots+2^{v\xi_k}\,d\boldsymbol{\xi} \\ &= 2^{1-k}\sum_{j=1}^{k}\int_0^1 2^{vy}\frac{y^{j-1}(1-y)^{k-j}}{(j-1)!(k-j)!}\,dy \\ &= \frac{2^{1-k}(2^v-1)}{v(k-1)!} \le \frac{20}{2^{k-v}k!}. \end{aligned}$$

□

14. Divisors of shifted primes

Proof of Theorem 6. We first take care of some easy cases. When $z \le z_0(y)$, $H(x,y,z) \asymp \eta x$. By the Brun-Titchmarsh inequality for primes in arithmetic progressions (e.g. [30]),

$$H(x,y,z;P_\lambda) \ll \frac{x}{\log x}\sum_{y<d\le z}\frac{1}{\phi(d)}.$$

The theorem in this case follows from the asymptotic formula (5.1).

If $z \ge y^{1.001}$, then $H(x,y,z) \asymp x$ and $H(x,y,z;P_\lambda) \le \pi(x) \ll \frac{x}{\log x}$. The most difficult case is when $z_0(y) \le z \le y^{1.001}$. We follow the outline from Section 3, inserting a sieve estimate at the appropriate point.

Lemma 14.1. *Suppose* $10 \le y < z = e^\eta y$, $y \le \sqrt{x}$, $(\log y)^{-1/2} \le \eta \le \frac{\log y}{1000}$. *Then*

$$H(x,y,z;P_\lambda) \ll x \max_{y^{1/2}\le t\le x} \sum_{P^+(a)\le te^\eta} \frac{L(a;\eta)/\phi(a)}{\log^2\left(\frac{t}{a}+P^+(a)\right)\log\left(\frac{x^{0.49}}{a}+P^+(a)\right)}.$$

The proof of Lemma 14.1 is similar to Lemma 3.2, and will be given at the end of this section. By the proof of Lemma 3.3 and the Cauchy-Schwarz inequality, we obtain

$$H(x,y,z;P_\lambda) \ll x(1+\eta) \max_{y^{1/2}\le t\le x} S^*(t;\eta)^{1/2}\widetilde{S}(t;\eta)^{1/2}, \tag{14.1}$$

where

$$\widetilde{S}(t;\eta) = \sum_{a\in\mathscr{P}^*(e^\eta,te^\eta)} \frac{a}{\phi^2(a)} \frac{L(a;\eta)}{\log^2(x^{0.4}/a + P^+(a))}.$$

In the same way that (3.8) was proved, we obtain

$$\widetilde{S}(t;\eta) \ll \frac{\widetilde{T}(\eta, te^\eta, 1)}{\log^2 x} + \sum_{\substack{k\in\mathbb{Z},k\ge 1\\ e^\eta \le e^{e^{k-1}} \le te^\eta}} e^{-2k}\widetilde{T}(\sigma, e^{e^k}, x^{1/10}),$$

where

$$\widetilde{T}(\sigma,P,Q) = \sum_{\substack{a\in\mathscr{P}^*(e^\sigma,P)\\ a\ge Q}} \frac{a}{\phi^2(a)} L(a;\sigma).$$

Similarly, we define $\widetilde{T}_k(\sigma,P,Q)$. The bound given in Lemma 3.5 holds with $T_k(\sigma,P,Q)$ replaced by $\widetilde{T}_k(\sigma,P,Q)$. The only change in the proof is to define E_j so that

$$\sum_{p\in E_j} \left(\frac{p}{(p-1)^2}\right)^{1-\gamma} \le \log 2.$$

Therefore, Lemma 3.9 holds with $S^*(t;\eta)$ replaced by $\widetilde{S}(t;\eta)$, and the theorem follows from (14.1). $\square$

Proof of Lemma 14.1. Suppose $q\le x$, q is prime and $\tau(q+\lambda,y,z)\ge 1$. We distinguish two cases. First, assume $z \le \exp\{(\log x)^{9/10}\}$. Let $y<d\le z$, $d|(q+\lambda)$ and set $p=P^+(d)$. Write $q+\lambda = apb$, where $P^+(a)\le p\le P^-(b)$. The number of q with $ap>\sqrt{x}$ is, by Lemma 2.5, at most

$$\sum_{\substack{m>\sqrt{x}\\ P^+(m)\le z}} \frac{x}{m} \ll \frac{x\log x}{\log z} e^{-\frac{\log x}{8\log z}} \ll \frac{x}{\log^{10} z}.$$

Now suppose $ap\le\sqrt{x}$. Given a,p, we wish to count the number of $b\le x/(ap)$ with $P^-(b)\ge p$ and $abp-\lambda$ prime. By the arithmetic form of the large sieve inequality [30], the number of b is

$$\ll \frac{x}{\phi(ap)\log x\log p} \ll \frac{x}{\log x}\frac{1}{\phi(a)p\log p}.$$

As in the proof of Lemma 6.1, for each a we have

$$\sum_{\substack{\tau(a,y/p,z/p)\ge 1\\ p\ge P^+(a)}} \frac{1}{p\log p} \ll \frac{L(a;\eta)}{\log^2(y/a+P^+(a))}$$

and Lemma 14.1 follows in this case.

Now suppose $z > \exp\{(\log x)^{9/10}\}$. Put $\sigma=\eta$ and suppose x_1, x_2 satisfy the hypotheses of Lemma 6.1. Define $\mathscr{A}$ as in that lemma, and let $\mathscr{A}^*$ be the set of $n\in\mathscr{A}$ for which $n-\lambda$ is prime. By the argument leading to (6.2), and since $\log z \ge (\log x)^{9/10}$, we obtain

$$H(x_2,y,z;P_\lambda) - H(x_1,y,z;P_\lambda) \le |\mathscr{A}^*| + O\left(\frac{x_2-x_1}{\log^{3.6} x_1}\right).$$

Assume (6.3) and write $q+\lambda$ in the form (6.4). Given a and p, we wish to count the number of $b\in(\frac{x_1}{ap},\frac{x_2}{ap}]$ with $P^-(b)>p$, $b>p$, and $apb-\lambda$ prime. Here

$$\frac{x_2-x_1}{ap} \ge \frac{x_1}{ap\log^{10} z} \ge \max\left(p^{1/2}, \frac{x_1^{0.999}}{az_j}\right) := Q.$$

We apply the arithmetic form of the large sieve [30], eliminating those b with $b\equiv 0 \pmod{\varpi}$ for some prime $\varpi < p^{1/10}$ and those b with $apb-\lambda\equiv 0 \pmod{\varpi}$ for some prime $\varpi\le Q^{1/3}$. The number of remaining b is

$$\ll \frac{x_2-x_1}{\phi(ap)\log p\log Q} \ll \frac{x_2-x_1}{\phi(a)p\log p\log(x_1^{0.999}/(az_j)+y_j/a+P^+(a))}.$$

As in the proof of Lemma 6.1, the sum over p of $\frac{1}{p\log p}$ is

$$\ll \frac{L(a;\eta)}{\log^2(y_j/a+P^+(a))}.$$

Thus

$$H(x_2,y,z;P_\lambda) - H(x_1,y,z;P_\lambda) \le (x_2-x_1)\Bigg[O\left(\frac{1}{\log^{3.6} x}\right)$$
$$+\sum_{j=1}^{2}\sum_{P^+(a)\le z_j} \frac{L(a;\eta)}{\phi(a)\log^2(y_j/a+P^+(a))\log(x_1^{0.999}/az_j+y_j/a+P^+(a))}\Bigg].$$

Since $z_j \le y_j^{1.001} \le x^{0.001} y_j$, we have

$$\frac{x_1^{0.999}}{z_j} + y_j \ge x_1^{0.499}.$$

By considering the term $a=1$ alone, we see that the double sum exceeds

$$\eta/\log^3 y_1 \ge (\log x_1)^{-3.6}.$$

Lastly, we sum over short intervals $(x_1,x_2]\subseteq[x/\log^5 x, x]$. $\square$

Proof of Theorem 7. First, suppose $0 < \alpha < \beta < \frac{1}{2}$. Let N_p denote the number of primes $q \in (x/2, x]$ with $q \equiv -\lambda \pmod p$. For large x, a given number $q + \lambda$ can be divisible by at most $\lfloor 1/\alpha \rfloor + 1$ primes $p > x^\alpha$. Thus, using the Bombieri-Vinogradov theorem (see e.g., Chapter 28 of [4]), we have

$$\begin{aligned} H(x, x^\alpha, x^\beta; P_\lambda) - H(x/2, x^\alpha, x^\beta; P_\lambda) &\ge \frac{1}{\lfloor 1/\alpha \rfloor + 1} \sum_{x^\alpha < p \le x^\beta} N_p \\ &\gg_{\alpha,\lambda} \frac{x}{\log x} \sum_{x^\alpha < p \le x^\beta} \frac{1}{p-1} \\ &\gg_{\alpha,\beta,\lambda} \frac{x}{\log x}. \end{aligned} \tag{14.2}$$

The theorem follows when $0 \le a < b \le \frac{1}{2}$ by applying (14.2) with α, β chosen so that $a < \alpha < \beta < b$. When $0 \le a < \frac{1}{2} < b$, the theorem follows by applying (14.2) with $a < \alpha < \frac{1}{2}$ and $\beta = \frac{1}{2}$. Finally, suppose $\frac{1}{2} \le a < b \le 1$ and let a', b' be numbers with $a < a' < b' < b$. For large x, $q \in (x/2, x]$ implies $q + \lambda \in (x/4, 2x]$. Thus, if $d | (q + \lambda)$ and $x^{1-b'} < d \le x^{1-a'}$, then the complementary divisor $\frac{q+\lambda}{d}$ lies in $(\frac{1}{4} x^{a'}, 2x^{b'}]$. For large x, $\frac{q+\lambda}{d} \in (x^a, x^b]$ and thus the theorem follows by taking $\alpha = 1 - b'$ and $\beta = 1 - a'$ in (14.2). □

UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN, URBANA, IL
E-mail address: ford@math.uiuc.edu

REFERENCES

[1] R. C. BAKER, G. HARMAN, and J. PINTZ, The difference between consecutive primes. II, *Proc. London Math. Soc.* (3) **83** (2001), 532–562.

[2] A. S. BESICOVITCH, On the density of certain sequences of integers, *Math. Ann.* **110** (1934), 336–341.

[3] C. COBELI, K. FORD, and A. ZAHARESCU, The jumping champions of the Farey series, *Acta Arith.* **110** (2003), 259–274.

[4] H. DAVENPORT, *Multiplicative Number Theory*, third ed., *Graduate Texts in Mathematics*, **74**, Springer-Verlag, New York, 2000; revised and with a preface by Hugh L. Montgomery.

[5] P. ERDŐS, Note on the sequences of integers no one of which is divisible by any other, *J. London Math. Soc.* **10** (1935), 126–128.

[6] ———, A generalization of a theorem of Besicovitch, *J. London Math. Soc.* **11** (1936), 92–98.

[7] ———, Some remarks on number theory, *Riveon Lematematika* **9** (1955), 45–48 (Hebrew. English summary).

[8] P. ERDŐS, An asymptotic inequality in the theory of numbers, *Vestnik Leningrad Univ.* **15** (1960), 41–49 (Russian).

[9] P. ERDŐS and R. L. GRAHAM, Old and new problems and results in combinatorial number theory, *Monographies de L'Enseignement Mathématique* (*Monographs of L'Enseignement Mathématique*) **28**, Université de Genève L'Enseignement Mathématique, Geneva, 1980.

[10] P. Erdős and M. Ram Murty, On the order of a (mod p), *Number Theory* (Ottawa, ON, 1996), *CRM Proc. Lecture Notes* **19**, Amer. Math. Soc., Providence, RI, 1999, pp. 87–97.

[11] P. Erdős and G. Tenenbaum, Sur la structure de la suite des diviseurs d'un entier, *Ann. Inst. Fourier* (Grenoble) **31** (1981), no. 1, ix, 17–37.

[12] ———, Sur les diviseurs consécutifs d'un entier, *Bull. Soc. Math. France* **111** (1983), no. 2, 125–145.

[13] M. Filaseta and O. Trifonov, On gaps between squarefree numbers. II, *J. London Math. Soc.* **45** (1992), 215–221.

[14] K. Ford, Integers with a divisor in $(y, 2y]$, in *Anatomy of Integers* (J.-M. De Koninck, A. Granville, and Florian Luca, eds.), *CRM Proc. and Lecture Notes* **46**, 65–80, Amer. Math. Soc., Providence, RI, 2008.

[15] ———, Sharp probability estimates for generalized Smirnov statistics, *Monatshefte Math.* **153** (2008), 205–216.

[16] 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.* **133** (2005), no. 12, 3463–3468 (electronic).

[17] K. Ford and I. E. Shparlinski, On curves over finite fields with Jacobians with small exponent, *Internat. J. Number Theory* **4** (2008), 819–826.

[18] K. Ford and G. Tenenbaum, The distribution of integers with at least two divisors in a short interval, *Quart. J. Math. Oxford* **58** (2007), 187–201.

[19] H. Halberstam and H.-E. Richert, *Sieve Methods*, Academic Press (A subsidiary of Harcourt Brace Jovanovich, Publishers), New York, 1974, *London Mathematical Society Monographs*, No. 4.

[20] H. Halberstam and K. F. Roth, *Sequences*, Oxford University Clarendon Press, 1966.

[21] R. R. Hall, The propinquity of divisors, *J. London Math. Soc.* **19** (1979), 35–40.

[22] ———, *Sets of Multiples*, *Cambridge Tracts in Mathematics* **118**, Cambridge University Press, Cambridge, 1996.

[23] R. R. Hall and G. Tenenbaum, Sur la proximité des diviseurs, *Recent Progress in Analytic Number Theory* **1** (Durham, 1979), Academic Press, London, 1981, pp. 103–113.

[24] ———, *Divisors*, *Cambridge Tracts in Mathematics* **90**, Cambridge University Press, Cambridge, 1988.

[25] ———, The set of multiples of a short interval, *Number Theory* (New York, 1989/1990), Springer-Verlag, New York, 1991, pp. 119–128.

[26] D. R. Heath-Brown, Linear relations amongst sums of two squares, *Number Theory and Algebraic Geometry*, *London Math. Soc. Lecture Note Ser.* **303**, Cambridge Univ. Press, Cambridge, 2003, pp. 133–176.

[27] K.-H. Indlekofer and N. M. Timofeev, Divisors of shifted primes, *Publ. Math. Debrecen* **60** (2002), 307–345.

[28] H. Maier and G. Tenenbaum, On the set of divisors of an integer, *Invent. Math.* **76** (1984), 121–128.

[29] H. Maier and G. Tenenbaum, On the normal concentration of divisors, *J. London Math. Soc.* **31** (1985), 393–400.

[30] H. L. Montgomery and R. C. Vaughan, The large sieve, *Mathematika* **20** (1973), 119–134.

[31] K. K. Norton, On the number of restricted prime factors of an integer. I, *Illinois J. Math.* **20** (1976), 681–705.

[32] F. Pappalardi, On the order of finitely generated subgroups of $\mathbf{Q}^*$ (mod p) and divisors of $p-1$, *J. Number Theory* **57** (1996), 207–222.

[33] A. Raouj and G. Tenenbaum, Sur l'écart quadratique moyen des diviseurs d'un entier normal, *Math. Proc. Camb. Phil. Soc.* **126** (1999), 399–415.

[34] J. Riordan, *Combinatorial Identities*, John Wiley & Sons Inc., New York, 1968.

[35] R. Sitaramachandra Rao, On an error term of Landau, *Indian J. Pure Appl. Math.* **13** (1982), 882–885.

[36] N. V. Smirnov (Smirnoff), Sur les écarts de la courbe de distribution empirique, *Rec. Math. N.S.* **6(48)** (1939), 3–26, Russian, French summary.

[37] G. Tenenbaum, Sur deux fonctions de diviseurs, *J. London Math. Soc.* **14** (1976), 521–526; Corrigendum: *J. London Math. Soc.* **17** (1978), 212.

[38] ———, Sur la répartition des diviseurs, *Séminaire Delange-Pisot-Poitou*, 17e année (1975/76), Théorie des nombres: Fasc. 2, Exp. No. G14, Paris, 1977, p. 5 pp. Secrétariat Math.

[39] ———, Lois de répartition des diviseurs. II, *Acta Arith.* **38** (1980/81), 1–36.

[40] ———, Lois de répartition des diviseurs. III, *Acta Arith.* **39** (1981), 19–31.

[41] ———, Sur la probabilité qu'un entier possède un diviseur dans un intervalle donné, Seminar on number theory (Paris, 1981/1982), *Progr. Math.* **38**, Birkhäuser Boston, Boston, MA, 1983, pp. 303–312.

[42] ———, Sur la probabilité qu'un entier possède un diviseur dans un intervalle donné, *Compositio Math.* **51** (1984), 243–263.

[43] ———, Un problème de probabilité conditionnelle en arithmétique, *Acta Arith.* **49** (1987), 165–187.

[44] ———, *Introduction to Analytic and Probabilistic Number Theory*, *Cambridge Studies in Advanced Mathematics* **46**, Cambridge University Press, Cambridge, 1995, translated from the second French edition (1995) by C. B. Thomas.

[45] ———, Sur l'écart quadratique moyen des diviseurs d'un entier normal. II, *Math. Proc. Cambridge Philos. Soc.* **138** (2005), 1–8.

(Received January 20, 2004)