Michigan Math. J. 60 (2011), 199–214

On the Greatest Common Divisor of a Number and Its Sum of Divisors

Paul Pollack

1. Introduction

A natural number n is called *perfect* if $\sigma(n) = 2n$ and *multiply perfect* whenever $\sigma(n)$ is a multiple of n. In 1956, Erdős published improved upper bounds on the counting functions of the perfect and multiply perfect numbers [2]. These estimates were soon superseded by a theorem of Wirsing [15] (Theorem B in this paper), but Erdős's methods remain of interest because they are applicable to more general questions concerning the distribution of $\gcd(n, \sigma(n))$. Erdős [2] describes some applications of this type but omits the proofs. In this paper, we prove corrected versions of his results and establish some new results in the same direction.

For real numbers $x \geq 1$ and $A \geq 1$, put

$$G(x, A) := \#\{n \leq x : \gcd(n, \sigma(n)) > A\}.$$

Theorem 1.1. *Let $\beta > 0$. If $x > x_0(\beta)$ and $A > \exp((\log\log x)^{\beta})$, then $G(x, A) \leq x/A^c$, where $c = c(\beta) > 0$.*

This is (more or less) Theorem 3 of [2], except that Erdős assumes instead that $A > (\log x)^{\beta}$. After stating Theorem 3, Erdős claims that his result is best possible in that, if A grows slower than any power of $\log x$, then one does not save a fixed power of A in the estimate for $G(x, A)$. Theorem 1.1 shows that this assertion is incorrect. Our next result shows that Theorem 1.1 is best possible in the sense Erdős intended.

Theorem 1.2. *Let $\beta = \beta(x)$ be a positive real-valued function of x satisfying $\beta(x) \to 0$ as $x \to \infty$. Let $\varepsilon > 0$. If x is sufficiently large (depending on ε and the choice of function β) and $2 \leq A \leq \exp((\log\log x)^{\beta})$, then $G(x, A) \geq x/A^{\varepsilon}$.*

For large values of A, one may deduce a stronger upper bound on $G(x, A)$ than that of Theorem 1.1 from the following estimate for the mean of $\gcd(n, \sigma(n))$.

Theorem 1.3. *For each $x \geq 3$, we have*

$$\sum_{n \leq x} \gcd(n, \sigma(n)) \leq x^{1 + c_1/\sqrt{\log\log x}},$$

where c_1 is an absolute positive constant.

Received June 12, 2009. Revision received November 2, 2009.
The author is supported by NSF award DMS-0802970.

For example, it follows immediately from Theorem 1.3 that, if $A \geq x^\delta$ (for a fixed $\delta > 0$), then $G(x, A) \leq x/A^{1+o(1)}$ as $x \to \infty$. Our next result asserts that the analogous lower bound holds in a wide range of A.

THEOREM 1.4. *Fix $\varepsilon > 0$. Then $G(x, A) \geq x/A^{1+o(1)}$ as $x \to \infty$, uniformly for $2 \leq A \leq x^{1-\varepsilon}$.*

Theorems 1.3 and 1.4 have the following immediate consequence.

COROLLARY 1.5. *For each fixed $\delta \in (0, 1)$, we have $G(x, x^\delta) = x^{1-\delta+o(1)}$ as $x \to \infty$.*

NOTATION. For the most part we use standard notation of analytic number theory. For example, we write $\omega(n)$ for the number of distinct prime factors of n and $\mathrm{rad}(n)$ for the product of the distinct primes dividing n. We put $\log_1 x := \max\{\log x, 1\}$, and we define inductively $\log_k x = \max\{1, \log(\log_{k-1} x)\}$. We emphasize that when the letter c appears with a subscript, it always refers to an absolute positive constant.

2. Proof of Theorem 1.1

We require several preliminaries. Theorem A assembles results due to Kátai and Subbarao (see [9, Thm. 1]) and Erdős, Luca, and Pomerance (cf. [3, Thm. 8, Cor. 10]). See also [2, Thm. 4].

THEOREM A. *For all n outside of a set of asymptotic density zero,* $\gcd(n, \sigma(n))$ *is the largest divisor of n supported on the primes not exceeding* $\log\log n$.

For each real u, the set of n with $\gcd(n, \sigma(n)) > (\log\log n)^u$ *possesses an asymptotic density $g(u)$. The function $g(u)$ is continuous everywhere, is strictly decreasing on* $[0, \infty)$, *and satisfies* $g(0) = 1$ *and* $\lim_{u\to\infty} g(u) = 0$. *Explicitly, we have*

$$g(u) := e^{-\gamma} \int_u^\infty \rho(t)\, dt$$

for all $u > 0$, where γ is the Euler–Mascheroni constant and ρ is the Dickman–de Bruijn function.

The next lemma is proved by Erdős and Nicolas as [4, Thm. 2]—except for the statement concerning uniformity (which, however, is clear from their proof).

LEMMA 2.1. *For each fixed $c \in (0, 1]$, the number of $n \leq x$ with*

$$\omega(n) > c \frac{\log x}{\log_2 x}$$

is $x^{1-c+o(1)}$ as $x \to \infty$. Moreover, the convergence of the $o(1)$ term to zero is uniform if c is restricted to a compact subset of $(0, 1]$.

The next result is implicit in the proof of [2, Thm. 1]; for the convenience of the reader, we repeat the argument here.

LEMMA 2.2. *Let $\varepsilon > 0$. If $m > m_0(\varepsilon)$ is squarefree, then there is some divisor d of m with* $\gcd(d, \sigma(d)) = 1$ *and* $d \geq m^{1/2-\varepsilon}$.

Proof. By replacing m by $m/2$ if necessary, we may assume that m is odd. We now run the following algorithm. Put $d_0 = 1$ and $d_0' = m$. Having defined d_i and d_i' so that $d_i d_i' = m$ and $\gcd(d_i, \sigma(d_i)) = 1$, we proceed as follows. If there is a prime dividing d_i' that does not divide $\sigma(d_i)$, then let p be the largest such prime and set $d_{i+1} = d_i p$ and $d_{i+1}' = d_i'/p$. (If there is no such prime, terminate the algorithm.) Then $d_{i+1}d_{i+1}' = m$ and

$$\begin{aligned}\gcd(d_{i+1}, \sigma(d_{i+1})) &= \gcd(d_i p, \sigma(d_i)(p+1)) \\ &= \gcd(d_i, p+1),\end{aligned}$$

since $p \nmid \sigma(d_i)$ and $\gcd(d_i, \sigma(d_i)) = 1$. Since p is odd (by our assumption that m is odd), every prime factor q of $p+1$ is smaller than p. None of these q can divide d_i; indeed, if q divides d_i, then there must be some $j < i$ for which q is the largest prime divisor of d_j' not dividing $\sigma(d_j)$. But this is absurd: $q < p$, p is a divisor of d_j' (since d_i' divides d_j'), and $p \nmid \sigma(d_j)$ (since $d_j \mid d_i$ and $p \nmid \sigma(d_i)$). Thus $\gcd(d_i, p+1) = 1$ and so $\gcd(d_{i+1}, \sigma(d_{i+1})) = 1$.

At the end of this algorithm we have numbers d_k, d_k' with $d_k d_k' = m$ and $\gcd(d_k, \sigma(d_k)) = 1$. Moreover, d_k' must divide $\sigma(d_k)$, for otherwise we could continue the algorithm. Recalling the maximal order of the sum of divisors function (see [6, Thm. 323]), we find that

$$d_k \log_2 d_k \gg \sigma(d_k) \geq d_k', \quad \text{whence} \quad d_k^2 \log_2 d_k \gg d_k d_k' = m.$$

Hence $d_k \gg (m/\log_2 m)^{1/2}$, so in particular $d_k \geq m^{1/2-\varepsilon}$ for large m. So if we choose $d = d_k$, then we have the lemma. □

The next lemma is an easy consequence of the Brun–Titchmarsh inequality; for a proof see, for example, [8, Lemma 6].

LEMMA 2.3. *Let m be a positive integer. For all $x \geq 1$, we have*

$$\sum_{\substack{p \leq x \\ p \equiv -1 \pmod{m}}} \frac{1}{p} \ll \frac{\log_2 x}{\varphi(m)}.$$

Here the implied constant is absolute.

LEMMA 2.4. *Let d be a squarefree integer. For $x \geq 1$, the number of squarefree $n \leq x$ for which for which d divides $\sigma(n)$ is at most*

$$\omega(d)^{\omega(d)} \frac{x}{\varphi(d)} (C \log_2 x)^{\omega(d)},$$

where C is an absolute positive constant.

Lemma 2.4 is similar to the case $k = 1$ of [1, Lemma 2]; however, the dependence on $\omega(d)$ is more explicit in our bound.

Proof of Lemma 2.4. Since d divides $\sigma(n) = \prod_{p|n}(p+1)$, we can write $d = \prod_{p|n} a_p$, where each a_p divides $p+1$. Throwing away those $a_p = 1$, we see that n induces a (not necessarily unique) factorization of d. By a *factorization* of d we mean a decomposition of d as a product of factors strictly larger than 1, where the order of the factors is not taken into account. For each possible factorization of d, we estimate the number of $n \le x$ (as in the lemma statement) that induce this factorization.

Let $d = a_1 a_2 \cdots a_k$ be a factorization of d. If n induces this factorization, then there are distinct primes $p_1, \dots, p_k$ dividing n with $p_i \equiv -1 \pmod{a_i}$ for each $1 \le i \le k$. So by Lemma 2.3, if C is an appropriate absolute positive constant then the number of such $n \le x$ is

$$\begin{aligned} &\le \sum_{p_1 \equiv -1 \,(\mathrm{mod}\, a_1)} \cdots \sum_{p_k \equiv -1 \,(\mathrm{mod}\, a_k)} \frac{x}{p_1 \cdots p_k} \\ &\le x \prod_{i=1}^{k} \frac{C \log_2 x}{\varphi(a_i)} = \frac{x}{\varphi(d)} (C \log_2 x)^k \le \frac{x}{\varphi(d)} (C \log_2 x)^{\omega(d)}. \end{aligned}$$

(The last inequality uses the observation that each factorization of d involves at most $\omega(d)$ factors.) Since d is squarefree, the number of factorizations of d is given by $B_{\omega(d)}$, where B_l (the lth *Bell number*) stands for the number of set partitions of an l-element set.

Since any partition of an l-element set involves at most l components, we always have $B_l \le l^l$. Taking $l = \omega(d)$ completes the proof of Lemma 2.4. $\square$

The last part of our preparation is reducing the proof of Theorem 1.1 to that of the following squarefree version.

PROPOSITION 2.5. *Let $\beta > 0$. If $x > x_1(\beta)$ and $A > \exp((\log_2 x)^{\beta})$, then the number of squarefree $n \le x$ with $\gcd(n, \sigma(n)) > A$ is at most $x/A^{c'}$, where $c' = c'(\beta)$.*

LEMMA 2.6. *Theorem 1.1 follows from Proposition 2.5.*

Proof. Let $\beta > 0$. Suppose that $n \le x$ and $\gcd(n, \sigma(n)) > A$, where $A > \exp((\log_2 x)^{\beta})$. Write $n = n_0 n_1$, where n_0 is squarefree, n_1 is squarefull, and $\gcd(n_0, n_1) = 1$. If $n_1 > A^{1/4}$, then n belongs to a set of size at most

$$x \sum_{\substack{m > A^{1/4} \\ m \text{ squarefull}}} m^{-1} \ll \frac{x}{A^{1/8}}.$$

(Here we use that the counting function of the squarefull integers is $O(\sqrt{x})$.) Otherwise, since

$$\begin{aligned} A &< \gcd(n_0 n_1, \sigma(n_0)\sigma(n_1)) \\ &\le \gcd(n_0, \sigma(n_0)) \gcd(n_0, \sigma(n_1)) \gcd(n_1, \sigma(n_0)\sigma(n_1)) \\ &\le \gcd(n_0, \sigma(n_0)) n_1 \sigma(n_1) \le \gcd(n_0, \sigma(n_0)) n_1^3, \end{aligned}$$

it follows that

$$\gcd(n_0, \sigma(n_0)) \geq A/n_1^3 \geq A^{1/4}.$$

The number of such $n \leq x$ is therefore at most

$$\sum_{\substack{n_0 \leq x \\ n_0 \text{ squarefree} \\ \gcd(n_0,\sigma(n_0))>A^{1/4}}} \sum_{\substack{n_1 \leq x/n_0 \\ n_1 \text{ squarefull} \\ \gcd(n_0,n_1)=1}} 1 \ll \sqrt{x} \sum_{\substack{n_0 \leq x \\ n_0 \text{ squarefree} \\ \gcd(n_0,\sigma(n_0))>A^{1/4}}} \frac{1}{\sqrt{n_0}}. \tag{1}$$

Define

$$B(u) := \sum_{\substack{m \leq u \\ m \text{ squarefree} \\ \gcd(m,\sigma(m))>A^{1/4}}} 1.$$

Since $A^{1/4} > \exp\left(\frac{1}{4}(\log_2 x)^{\beta}\right) > \exp((\log_2 x)^{\beta/2})$ for large x, we can apply Proposition 2.5 with β replaced by $\beta/2$ to find that $B(u) \leq u/A^{c'/4}$, where $c' = c'(\beta/2)$ and the inequality holds for all $u \leq x$ that are large enough (depending only on β). Partial summation now shows that, for sufficiently large x (depending just on β), the final sum in (1) is $\ll x^{1/2}/A^{c'/4}$; hence the double sum in (1) is $\ll x/A^{c'/4}$.

It follows that Theorem 1.1 holds if $c = c(\beta)$ is chosen as any constant smaller than $\min\left\{\frac{1}{8}, \frac{c'}{4}\right\}$. □

We now prove Proposition 2.5 (and so also Theorem 1.1). Assume that $\beta > 0$, $A > \exp((\log_2 x)^{\beta})$, and n is a squarefree integer with $\gcd(n, \sigma(n)) > A$. Write D for $\gcd(n, \sigma(n))$.

If there is a prime $p > A^{1/2}$ dividing D, then n has the form pr, where $p \mid \sigma(r)$. By Lemma 2.4, the number of possible r is

$$\ll \frac{x/p}{\varphi(p)} \log_2 x \ll \frac{x \log_2 x}{p^2},$$

so that the number of n that can arise in this way is

$$\ll x \log_2 x \sum_{p > A^{1/2}} \frac{1}{p^2} \ll \frac{x \log_2 x}{A^{1/2}}.$$

This number is smaller than $x/A^{1/3}$ for large x (depending on β).

We may therefore assume that the largest prime dividing D is at most $A^{1/2}$. Since $D > A$, by successively stripping primes from D we must eventually discover a divisor of D in the interval $(A^{1/2}, A]$. If x (and hence A) is large, we can apply Lemma 2.2 (with $\varepsilon = 1/6$) to this divisor to obtain a divisor d of D with $d \in (A^{1/6}, A]$ having the property that $\gcd(d, \sigma(d)) = 1$.

Write $n = de$. Since $d \mid \sigma(n)$ and $\gcd(d, \sigma(d)) = 1$, it follows that $e \leq x/d$ and $d \mid \sigma(e)$. By Lemma 2.4, the number of such e is

$$\leq \frac{x}{d\varphi(d)} (C\omega(d) \log_2 x)^{\omega(d)}. \tag{2}$$

The strategy for the rest of the proof is as follows. First, if $\omega(d)$ is not too large, then (2) is manageable and summing over such d yields an acceptable bound on the number of corresponding n. Otherwise, n is divisible by some $d \in (A^{1/6}, A]$ with an abnormally large number of prime divisors, and Lemma 2.1 implies that such n are rare.

Let c be a small constant (depending only on β) whose value will be chosen momentarily. Suppose that

$$\omega(d) < c\frac{\log A}{\log\log A}.$$

Then, for large x,

$$(C\omega(d)\log_2 x)^{\omega(d)} \leq \exp\left(c\frac{\log A}{\log_2 A}(\log_2 A + \log_3 x)\right).$$

Since $A > \exp((\log_2 x)^\beta)$, we have $\log_2 A > \beta \log_3 x$, so this upper bound is at most

$$\exp(c(1+\beta^{-1})\log A) = A^{c(1+\beta^{-1})}.$$

We now assume $c > 0$ is small enough that $c(1+\beta^{-1}) \leq 1/12$. Then, by summing (2) over these values of d, we obtain an upper bound on the number of corresponding n that is at most

$$xA^{1/12}\sum_{d>A^{1/6}}\frac{1}{d\varphi(d)} \ll \frac{x}{A^{1/12}},$$

using Landau's result that $\sum_{d\geq t}\frac{1}{d\varphi(d)} \ll \frac{1}{t}$.

The remaining n have a divisor $d \in (A^{1/6}, A]$ for which $\omega(d) > c\log A/\log\log A$; the number of such n is at most $x\sum\frac{1}{d}$, where the sum is extended over all such d. Let

$$B(u) := \sum_{\substack{m\leq u\\ \omega(m)>c\log A/\log_2 A}} 1.$$

For $A^{1/6} \leq u \leq A$, define d_u as follows:

$$d_u\frac{\log u}{\log_2 u} = \frac{\log A}{\log_2 A} \quad \text{so that } d_u = (1+o(1))\frac{\log A}{\log u} \quad \text{as } u \to \infty.$$

By Lemma 2.1, for these u we have

$$B(u) = u^{1-cd_u+o(1)} = u^{1-c\log A/\log u}A^{o(1)} = (u/A^c)A^{o(1)} \leq u/A^{c/2},$$

say. (Note that, for large x, the real number cd_u belongs to the compact subinterval $[c/2, 12c]$ of $(0,1]$.) Thus

$$\sum_{\substack{d\in(A^{1/6},A]\\ \omega(d)>c\log A/\log_2 A}}\frac{1}{d} = \frac{B(A)}{A} - \frac{B(A^{1/6})}{A^{1/6}} + \int_{A^{1/6}}^{A}\frac{B(t)}{t^2}\,dt$$
$$\ll A^{-c/2} + (\log A)A^{-c/2} \ll A^{-c/3},$$

say.

Piecing everything together, it follows that the number of $n \leq x$ with $\gcd(n, \sigma(n)) > A$ is at most $x/A^{c'(\beta)}$, for large x, if we choose $c'(\beta) < \min\left\{\frac{1}{3}c, \frac{1}{12}\right\}$.

3. Proof of Theorem 1.2

We begin by recalling some results from the theory of smooth numbers. Let $\Psi(x, y)$ denote the number of y-smooth positive integers $n \le x$, where n is called *y-smooth* if each prime p dividing n satisfies $p \le y$. Let $\Psi_2(x, y)$ denote the number of squarefree y-smooth numbers $n \le x$. The following estimate of de Brujin appears as [14, Thm. 2, p. 359].

LEMMA 3.1. *Uniformly for $x \ge y \ge 2$,*

$$\log \Psi(x, y) = Z\left(1 + O\left(\frac{1}{\log y} + \frac{1}{\log_2 2x}\right)\right),$$

where

$$Z := \frac{\log x}{\log y} \log\left(1 + \frac{y}{\log x}\right) + \frac{y}{\log y} \log\left(1 + \frac{\log x}{y}\right).$$

The following result is due to Ivić and Tenenbaum [7] and Naimi [12] (independently).

LEMMA 3.2. *Whenever $x, y \to \infty$ and $\log y/\log_2 x \to \infty$, we have $\Psi_2(x, y) = (6/\pi^2 + o(1))\Psi(x, y)$.*

The next lemma is due to Pomerance (cf. [13, Thm. 2]).

LEMMA 3.3. *Let $x \ge 3$ and let m be a positive integer. The number of $n \le x$ for which $m \nmid \sigma(n)$ is $\ll x/(\log x)^{1/\varphi(m)}$, where the implied constant is absolute.*

We now have all the tools at our disposal necessary to prove Theorem 1.2. By Theorem A we may assume that

$$\log_2 x < A < \exp((\log_2 x)^{\beta(x)}). \tag{3}$$

Put $y := (\log_2 x)^{1-\sqrt{\beta(x)}}$.

LEMMA 3.4. *If x is sufficiently large (depending on the choice of the function β), then all but at most x/A numbers $n \le x$ are such that $\sigma(n)$ is divisible by every prime $p \le y$.*

Proof. By Lemma 3.3, the number of exceptional n is

$$\ll y \frac{x}{(\log x)^{1/y}} \le (\log_2 x) \frac{x}{\exp((\log_2 x)^{\sqrt{\beta}})}.$$

To see that this is at most x/A, note that by the upper bound on A in (3) and a short computation, it is enough to prove that

$$\log_3 x - (\log_2 x)^{\sqrt{\beta}} < -(\log_2 x)^{\beta}.$$

From (3) we have that $(\log_2 x)^{\beta} > \log_3 x$ and so, for large x,

$$\begin{aligned}(\log_2 x)^{\sqrt{\beta}} - (\log_2 x)^{\beta} &= ((\log_2 x)^{\beta})^{1/\sqrt{\beta}} - (\log_2 x)^{\beta} \\ &> ((\log_2 x)^{\beta})^2 - (\log_2 x)^{\beta} \\ &> (\log_3 x)^2 - \log_3 x > \log_3 x,\end{aligned}$$

which gives the lemma. □

LEMMA 3.5. *If x is sufficiently large (depending on β and ε), then the number of positive integers $n \le x$ that have a squarefree, y-smooth divisor in the interval $(A, A^2]$ is at least $x/A^{\varepsilon/2}$.*

Proof. Let $P_y := \prod_{p \le y} p$ be the product of the primes not exceeding y. The number of $n \le x$ with a squarefree, y-smooth divisor $d \in (A, A^2]$ is

$$\ge \sum_{\substack{d \mid P_y \\ A < d \le A^2}} \sum_{\substack{n \le x \\ d \mid n,\ (n/d, P_y) = 1}} 1. \tag{4}$$

By inclusion-exclusion and Mertens's theorem, for each d in the outer sum, the inner sum is

$$\left(\frac{x}{d}\right)\frac{e^{-\gamma}}{\log y} + O(2^{\log_2 x}) = (e^{-\gamma} + o(1))\frac{x}{d \log_3 x} \quad (\text{as } x \to \infty),$$

so the double sum (4) is

$$\gg \frac{x}{\log_3 x} \sum_{\substack{d \mid P_y \\ A < d \le A^2}} \frac{1}{d} \ge \frac{x}{\log_3 x} \frac{1}{A^2} (\Psi_2(A^2, y) - A). \tag{5}$$

As $x \to \infty$, we have

$$\frac{\log y}{\log_2(A^2)} \ge (1 + o(1)) \frac{\log_3 x}{\beta(x) \log_3 x + \log 2},$$

and this lower bound tends to infinity with x as $\beta(x)$ tends to zero. So by Lemma 3.2 we have that $\Psi_2(A^2, y) \sim (6/\pi^2)\Psi(A^2, y)$. Since $\log(A^2) = y^{o(1)}$, Lemma 3.1 implies that (for $x \to \infty$)

$$\Psi(A^2, y) \ge \exp((1 + o(1)) \log(A^2)) = A^{2+o(1)}.$$

Referring back to (5), we find that the double sum (4) is bounded below by $(x/\log_3 x)A^{o(1)}$, which is at least $xA^{o(1)}$ because $A \ge \log_2 x$.

Theorem 1.2 follows immediately from Lemmas 3.4 and 3.5. Indeed, with at most x/A exceptions, any n with a divisor of the form prescribed in Lemma 3.5 will satisfy $\gcd(n, \sigma(n)) > A$. Since there are at least

$$x/A^{\varepsilon/2} - x/A > x/A^{\varepsilon}$$

such n, we have Theorem 1.2.

4. Proof of Theorem 1.3

For each natural number n, write $\sigma(n)/n = a(n)/b(n)$, where $a(n)$ and $b(n)$ are coprime natural numbers. Thus $a(n) = \sigma(n)/\gcd(n,\sigma(n))$ and $b(n) = n/\gcd(n,\sigma(n))$. The proof of Theorem 1.3 rests on the following theorem of Wirsing [15].

THEOREM B. *For each $x \geq 1$ and every pair of positive integers a and b, the number of $n \leq x$ for which $a(n) = a$ and $b(n) = b$ is*

$$\leq x^{c_2/\log_2 x}.$$

Here c_2 denotes an absolute positive constant.

For our purposes, we require a variant of Theorem A in which only the denominator is specified. It is perhaps surprising that a useful result of this kind can be derived from Theorem A by a simple inductive argument.

THEOREM 4.1. *For each $x \geq 1$ and each positive integer b, the number of $n \leq x$ for which $b(n) = b$ is*

$$\leq x^{c_3/\sqrt{\log_2 x}}.$$

Here c_3 is an absolute positive constant.

Let us suppose temporarily that Theorem 4.1 has been established. Then we can quickly dispense with Theorem 1.3 following a method of Erdős, Luca, and Pomerance (cf. the proof of the upper bound in [3, Thm. 11]). Indeed, for $x \geq 1$,

$$\frac{1}{x}\sum_{n\leq x}\gcd(n,\sigma(n)) \leq \sum_{n\leq x}\frac{\gcd(n,\sigma(n))}{n} = \sum_{b\leq x}\frac{1}{b}\sum_{\substack{n\leq x\\ b(n)=b}} 1$$
$$\leq (1+\log x)x^{c_3/\sqrt{\log_2 x}} \leq x^{c_1/\sqrt{\log_2 x}}$$

for an appropriate constant c_1. So it is enough to prove Theorem 4.1.

LEMMA 4.2. *Suppose that $x \geq 1$. For each positive integer $b \leq x$, the number of $n \leq x$ with $\mathrm{rad}(n) \mid b$ is at most $x^{c_4/\log_2 x}$.*

We remark that Lemma 4.2 is implicit in the proof of [3, Thm. 11].

Proof of Lemma 4.2. For a given x, the number of such n is maximized when b is the largest product of consecutive primes (starting at 2) not exceeding x. In this case the number of such n is precisely $\Psi(x,p)$, where p is the largest prime divisor of b. By the prime number theorem, $p \sim \log x$, and by Lemma 3.1, $\Psi(x,p) = x^{(\log 4+o(1))/\log_2 x}$ as $x \to \infty$. $\square$

Proof of Theorem 4.1. By [6, Thm. 323] we can fix a real number $x_0 > e^{2e}$ with the property that, for all $x \geq x_0$, we have

$$\sigma(n)/n \le 2\log_2 x$$

whenever $n \le x$. We prove that for each integer $N \ge 2$, each $x > x_0^{N/2}$, and each positive integer b, the number of $n \le x$ for which $b(n) = b$ is bounded by

$$x^{1/N + c_5 N/\log_2 x}. \tag{6}$$

Theorem 4.1 follows for large x upon choosing $N = \lfloor \sqrt{\log_2 x} \rfloor$. This implies the same theorem for all $x \ge 3$ with a possibly different constant in the exponent.

We proceed by induction on N. Suppose first that $N = 2$. If $b(n) = b$, then b divides n and so we can assume $b \le x^{1/2}$ (since otherwise we obtain a bound $x^{1/2}$ on the number of possible n, which is sharper than (6) for $N = 2$). Since $x > x_0$ by hypothesis, every $n \le x$ with $b(n) = b$ has

$$a(n) \le 2b\log_2 x \le 2x^{1/2}\log_2 x.$$

So by Wirsing's theorem (Theorem B), we know that the number of such n is at most

$$2x^{1/2}(\log_2 x)x^{c_2/\log_2 x} \le x^{1/2}x^{2c_5/\log_2 x}$$

if c_5 is chosen appropriately (depending on x_0 and c_2). This proves the upper bound (6) for $N = 2$.

Suppose the bound (6) is known for N; we prove it holds also for $N+1$. If $b \le x^{1/(N+1)}$, then we can apply Wirsing's theorem as before to obtain that the number of $n \le x$ with $b(n) = b$ is bounded by

$$\begin{aligned} 2x^{1/(N+1)}(\log_2 x)x^{c_2/\log_2 x} &\le x^{1/(N+1)}x^{2c_5/\log_2 x} \\ &< x^{1/(N+1)}x^{(N+1)c_5/\log_2 x}, \end{aligned}$$

yielding (6). So we may suppose $b > x^{1/(N+1)}$. We can also assume $b \le x$, since otherwise there are no solutions $n \le x$ to $b(n) = b$. Suppose n is a solution to $b(n) = b$, and let d denote the largest divisor of n supported on the primes dividing b. Since $b \mid n$, we have $b \mid d$. Moreover, defining n' by the equation $n = dn'$, we have

$$n' = n/d \le x/b \le x^{N/(N+1)}$$

and (since $\gcd(d, n') = 1$)

$$\frac{\sigma(n')}{n'} = \frac{d}{\sigma(d)}\frac{\sigma(n)}{n} = \frac{d}{\sigma(d)}\frac{a}{b},$$

where $a = a(n)$. It follows that $b(n')$ divides $\sigma(d)b$. Let b' be an arbitrary divisor of $\sigma(d)b$. Since $x > x_0^{(N+1)/2}$ by hypothesis, it follows that

$$x^{N/(N+1)} \ge (x_0^{(N+1)/2})^{N/(N+1)} = x_0^{N/2}.$$

Hence, by the induction hypothesis, for each choice of b' there are at most

$$(x^{N/(N+1)})^{1/N}x^{c_5 N/\log_2 x} = x^{1/(N+1)}x^{c_5 N/\log_2 x}$$

possibilities for $n' \le x^{N/(N+1)}$ with $b(n') = b'$. (Here we have also used that $x^{N/(N+1)} > e^e$ and that the function $t^{1/\log_2 t}$ is increasing for $t > e^e$.) The maximal order of the divisor function (see e.g. [6, Thm. 317]) guarantees that the number of choices for b', given d, is bounded by $x^{c_6/\log_2 x}$; whereas, by Lemma 4.2, the number of choices for d is bounded by $x^{c_4/\log_2 x}$. It follows that the number of choices for $n = dn'$ is at most

$$x^{1/(N+1)} x^{(c_5 N + (c_6 + c_4))/\log_2 x} \leq x^{1/(N+1)} x^{c_5(N+1)/\log_2 x}$$

if we choose c_5 such that $c_5 \geq c_6 + c_4$. □

REMARK. Suppose $f : \mathbf{N} \to \mathbf{N}$ is a multiplicative function. Say that f has *property W* if the following holds (for each $\varepsilon > 0$):

> For $x > x_0(\varepsilon)$, the number of $n \leq x$ with $f(n)/n = a/b$ is bounded by x^ε, uniformly in the choice of positive integers a and b.

Say that f has *property W'* if the following holds (for each $\varepsilon > 0$):

> For $x > x_1(\varepsilon)$, the number of $n \leq x$ for which n divides $bf(n)$ is bounded by x^ε, uniformly for positive integers $b \leq x$.

Wirsing's argument establishes that property W holds for a large class of multiplicative functions (see [11] for a general statement as well an extension to certain compositions of multiplicative functions). The proof of Theorem 4.1 shows that if f has property W and $f(n) \ll_\rho n^{1+\rho}$ for each $\rho > 0$, then f also has property W'.

5. Proof of Theorem 1.4

It is convenient to divide the proof of Theorem 1.4 into two parts depending on the size of A. Clearly the theorem is entailed by the following two propositions.

PROPOSITION 5.1. *There is an absolute constant $c_7 > 0$ for which the following statement holds: for each fixed $\varepsilon > 0$ and each $A \in [2, x^{c_7}]$,*

$$G(x, A) > x/A^{1+\varepsilon}$$

if x is sufficiently large (depending only on ε).

PROPOSITION 5.2. *Fix $\varepsilon > 0$. We have $G(x, A) \geq x/A^{1+o(1)}$ as $x \to \infty$, uniformly for $x^\varepsilon \leq A \leq x^{1-\varepsilon}$.*

5.1. Proof of Proposition 5.1

LEMMA 5.3. *For some c_7, the following holds. For each fixed $\varepsilon > 0$, all large enough x (depending on ε), and all A with*

$$\log_2 x \leq A \leq x^{c_7},$$

there are at least $x/A^{1+2\varepsilon}$ positive integers $m \leq x/A^{1+\varepsilon}$ possessing both of the following properties:

(1) *there is a prime $q \parallel m$ for which $q+1$ has a prime divisor in $(A, A^{1+\varepsilon}]$;*
(2) *the least prime divisor of m exceeds $A^{1+\varepsilon}$.*

Proof of Proposition 5.1 assuming Lemma 5.3. We can assume $A \geq \log_2 x$, since for smaller values of A the result of Proposition 5.1 follows from Theorem A.

We now apply Lemma 5.3. For each $m \leq x/A^{1+\varepsilon}$ satisfying (1) and (2), choose a prime $q \parallel m$ for which $q+1$ has a prime divisor $p \in (A, A^{1+\varepsilon}]$ and form the number $n = mp$. Then $n \leq x$, $p \mid n$, and $p \mid q+1 \mid \sigma(n)$, so that $\gcd(n, \sigma(n)) \geq p > A$. Moreover, condition (2) guarantees that the numbers n formed in this way are all distinct. Thus, there are at least $x/A^{1+2\varepsilon}$ values of $n \leq x$ for which $\gcd(n, \sigma(n)) > A$. Replacing ε by $\varepsilon/2$ then yields the proposition. □

Proof of Lemma 5.3. The proof proceeds in several stages. We can (and do) assume $0 < \varepsilon < 1$. Put

$$\mathcal{Q} := \{q \text{ prime} : q+1 \text{ has a prime divisor in } (A, A^{1+\varepsilon}]\},$$

and let $Q(y) := \#(\mathcal{Q} \cap [1, y])$. We estimate the number of $q \leq y$ that are *not* in $\mathcal{Q}$. The fundamental lemma of the sieve (see e.g. [5, Thm. 2.5′]) provides an estimate of this quantity if we assume y exceeds a large fixed power of A. Indeed, supposing that $y \geq A^M$, where M is large but fixed, we obtain that the number of such q is (for x sufficiently large)

$$\leq (1+\delta_M)\left(\prod_{A < p \leq A^{1+\varepsilon}} \frac{p-2}{p-1}\right) \mathrm{Li}(y).$$

Here δ_M is a constant that tends to zero as $M \to \infty$ (and Li is the usual logarithmic integral).

Since ε is fixed, the product here tends to a constant $B(\varepsilon) < 1$ as x (and hence A) tends to infinity. We now fix an $M > 2$ with $(1+\delta_M)B(\varepsilon) < 1$. For this choice of M, the proportion of primes $\leq y$ not in $\mathcal{Q}$ is strictly less than 1. We have thus arranged that $Q(y) \gg \mathrm{Li}(y)$ whenever $y \geq A^M$ and x is sufficiently large.

We use this lower bound to estimate the number of $m \leq x/A^{1+\varepsilon}$ having properties (1) and (2) of the lemma. By the fundamental lemma of the sieve (this time see [5, Thm. 2.5]), if $A \leq x^{c_7}$, where c_7 is a sufficiently small absolute constant, then the number of $m \leq x/A^{1+\varepsilon}$ having property (2) is

$$\geq \frac{1}{2} \frac{x}{A^{1+\varepsilon}} \prod_{p \leq A^{1+\varepsilon}} \left(1 - \frac{1}{p}\right) \tag{7}$$

for large x.

If we require that m not only have property (2) but also have no prime divisors in $\mathcal{Q}$ exceeding A^M, then the number of such $m \leq x/A^{1+\varepsilon}$ is (by [5, Thm. 2.2])

$$\ll \frac{x}{A^{1+\varepsilon}} \prod_{p \leq A^{1+\varepsilon}} \left(1 - \frac{1}{p}\right) \prod_{\substack{A^M \leq q \leq x/A^{1+\varepsilon} \\ q \in \mathcal{Q}}} \left(1 - \frac{1}{q}\right),$$

where the implied constant is absolute. (Our previous assumption $M > 2$ ensures that the products here are over disjoint sets of primes.) The product over the primes in $\mathcal{Q}$ is

$$\leq \exp\Bigg(-\sum_{\substack{A^M \leq q \leq x/A^{1+\varepsilon} \\ q \in \mathcal{Q}}} \frac{1}{q}\Bigg).$$

Since $\mathcal{Q}(y) \gg \mathrm{Li}(y)$ for $y \geq A^M$, partial summation shows that the sum inside the exponential can be made arbitrarily large by choosing c_7 sufficiently small (perhaps depending on M). So, for a suitable choice of c_7, we obtain a bound of

$$\frac{1}{4} \frac{x}{A^{1+\varepsilon}} \prod_{p \leq A^{1+\varepsilon}} \left(1 - \frac{1}{p}\right)$$

on the number of such m.

Comparing this with (7), we see that there are at least

$$\frac{1}{4} \frac{x}{A^{1+\varepsilon}} \prod_{p \leq A^{1+\varepsilon}} \left(1 - \frac{1}{p}\right) \gg \frac{x}{A^{1+\varepsilon} \log A} \tag{8}$$

positive integers $m \leq x/A^{1+\varepsilon}$ with property (2) that also have some prime divisor in $\mathcal{Q}$ exceeding A^M. But such an m has property (1) of the lemma unless m is divisible by q^2 for some $q \in \mathcal{Q}$ exceeding A^M, and the number of such m is

$$\ll \frac{x}{A^{1+\varepsilon}} \sum_{q > A^M} \frac{1}{q^2} \ll \frac{x}{A^{1+\varepsilon+M}},$$

which is negligible compared to the lower bound (8).

So, there are $\gg x/(A^{1+\varepsilon} \log A)$ values of $m \leq x/A^{1+\varepsilon}$ with both properties (1) and (2), which suffices to yield the claim. $\square$

5.2. Proof of Proposition 5.2

The proof of Proposition 5.2 is based on entirely different principles from that of Proposition 5.1.

Let ψ denote the *Dedekind ψ-function,* which is the arithmetic function defined by $\psi(n) := n \prod_{p|n}(1 + 1/p)$. (Thus $\psi \leq \sigma$ pointwise, and ψ and σ agree on squarefree arguments.) For each integer $K \geq 0$, define

$$F_K(n) := \prod_{0 \leq k \leq K} \psi_k(n),$$

where ψ_k denotes the kth iterate of ψ. We need the following lemma.

LEMMA 5.4. *Let K be a fixed nonnegative integer. For each positive integer n, write*

$$F_K(n) = MN, \quad \textit{where } M := \prod_{\substack{p^e \| F_K(n) \\ p \leq \log^3 x}} p^e \textit{ and } N := \prod_{\substack{p^e \| F_K(n) \\ p > \log^3 x}} p^e.$$

Then, as $x \to \infty$, for all but $o(x)$ values of $n \leq x$ we have that N is squarefree and

$$M \leq \exp(2(5 \log_2 x)^{K+2}) = x^{o(1)}.$$

Luca and Pomerance [10, Sec. 3.2] showed the analogous result with the Euler φ-function replacing ψ. Since the proof of Lemma 5.4 is essentially identical to their arguments, we omit it.

Put $R_K(n) := \mathrm{rad}(F_K(n))$.

LEMMA 5.5. *Let K be a fixed positive integer. As $x \to \infty$, for all but $o(x)$ values of $n \in [x/2, x]$ we have*
$$R_K(n) = x^{K+1+o(1)}$$
and
$$\gcd(R_K(n), \psi(R_K(n))) > x^{K+o(1)}.$$

Proof. For all but $o(x)$ values of $n \in [x/2, x]$, the conclusion of Lemma 5.4 holds. For these typical n, we have
$$R_K(n) \geq \frac{F_K(n)}{M} \geq \frac{n^{K+1}}{M} \geq \frac{1}{2^{K+1}M} x^{K+1} = x^{K+1+o(1)}$$
and
$$R_K(n) \leq F_K(n) \leq x^{K+1}(2\log_2 x)^{1+2+\cdots+K} \leq x^{K+1+o(1)}.$$
(Here we use once again the maximal order of the sigma function.) This gives the first assertion of the lemma.

Moreover, in the notation of Lemma 5.4, N divides $R_K(n)$ for these n, so that $\psi(N)$ divides $\psi(R_K(n))$ and hence $\gcd(R_K(n), \psi(R_K(n))) \geq \gcd(N, \psi(N))$. Thus it is enough to show that, for these n, we have $\gcd(N, \psi(N)) \geq x^{K+o(1)}$.

For a positive integer m, define $\mathrm{rad}'(m)$ to be the product of the distinct primes dividing m that exceed $\log^3 x$. Since N is squarefree, it follows that
$$N = \mathrm{rad}'(F_K(n)) = \prod_{k=0}^{K} \mathrm{rad}'(\psi_k(n)).$$
The $K+1$ factors in the right-hand product are pairwise coprime, so
$$\begin{aligned}\gcd(N, \psi(N)) &= \prod_{k=0}^{K} \gcd(\mathrm{rad}'(\psi_k(n)), \psi(N)) \\ &\geq \prod_{k=1}^{K} \gcd\big(\mathrm{rad}'(\psi_k(n)), \psi(\mathrm{rad}'(\psi_{k-1}(n)))\big).\end{aligned}$$
Now we observe that
$$\mathrm{rad}'(\psi_k(n)) \mid \psi(\mathrm{rad}'(\psi_{k-1}(n))).$$
Indeed, suppose p divides $\psi_k(n)$ and $p > \log^3 x$. Then either p^2 divides $\psi_{k-1}(n)$ or q does for some prime $q \equiv -1 \pmod p$. Since N is squarefree, only the latter is possible. Now q divides $\mathrm{rad}'(\psi_{k-1}(n))$ and so
$$p \mid q+1 = \psi(q) \mid \psi(\mathrm{rad}'(\psi_{k-1}(n))).$$
Hence,
$$\begin{aligned}\gcd(N, \psi(N)) &\geq \prod_{k=1}^{K} \mathrm{rad}'(\psi_k(n)) = \frac{N}{\mathrm{rad}'(\psi_0(n))} \\ &\geq \frac{N}{n} = \frac{F_K(n)}{Mn} \geq \frac{n^K}{M} \geq \frac{1}{2^K M} x^K = x^{K+o(1)}.\end{aligned}$$
This completes the proof of Lemma 5.5. □

Proof of Proposition 5.2. Fix a positive integer K large enough that $1/K < \varepsilon/2$, and let δ denote a fixed real number with

$$0 < \delta < \frac{\varepsilon}{2}(K+1)^{-1}. \tag{9}$$

Put

$$\mathcal{I} := \left[\frac{1}{2}A^{1/K+\delta}, A^{1/K+\delta}\right].$$

Then, by Lemma 5.5, for almost all $n \in \mathcal{I}$ we have (as $x \to \infty$)

$$R_K(n) \le A^{1+1/K+(K+1)\delta+o(1)} \tag{10}$$

and

$$\gcd(R_K(n), \sigma(R_K(n))) \ge (A^{1/K+\delta})^{K+o(1)} > A.$$

Note that, from our choice of K and the inequalities (9) and (10), for these typical n we have

$$R_K(n) \le A^{1+\varepsilon+o(1)} \le (x^{1-\varepsilon})^{1+\varepsilon+o(1)} \le x^{1-\varepsilon^2/2}. \tag{11}$$

We now let $\mathcal{R}$ be the set of values $R_K(n)$ that arise from these typical $n \in \mathcal{I}$. Since $\mathrm{rad}(n) \mid R_K(n)$, each element of $\mathcal{R}$ arises from at most $x^{o(1)}$ values of n (by Lemma 4.2) and hence

$$\#\mathcal{R} \ge A^{1/K+\delta}x^{o(1)} = A^{1/K+\delta+o(1)}.$$

(We have $x^{o(1)} = A^{o(1)}$ because, by hypothesis, $\log x \asymp \log A$.) For each $r \in \mathcal{R}$, define

$$\mathcal{A}(r) := \{br : 1 \le b \le x/r \text{ and } \gcd(b,r) = 1\} \quad \text{and} \quad \mathcal{A} := \bigcup_{r\in\mathcal{R}} \mathcal{A}(r).$$

Note that every element br of $\mathcal{A}(r)$ satisfies

$$\gcd(br, \sigma(br)) = \gcd(br, \sigma(b)\sigma(r)) \ge \gcd(r, \sigma(r)) > A.$$

So the proof will be complete if we establish a suitable lower bound on $\#\mathcal{A}$. Using (11) together with inclusion-exclusion, we see that

$$\#\mathcal{A}(r) = \frac{x}{r}\frac{\varphi(r)}{r} + O(2^{\omega(r)}) \ge \frac{x}{r}x^{o(1)}$$

since $2^{\omega(r)} = d(r) \ll x^{\varepsilon^2/4}$, say. Moreover, each element $a \in \mathcal{A}$ is contained in at most $d(a) = x^{o(1)}$ of the sets $\mathcal{A}(r)$. It follows that

$$\begin{aligned}\#\mathcal{A} &\ge x^{o(1)}\,\#\mathcal{R}\Big(\min_{r\in\mathcal{R}} \#\mathcal{A}(r)\Big)\\ &\ge x^{o(1)}(A^{1/K+\delta+o(1)})\frac{x}{A^{1+1/K+(K+1)\delta+o(1)}} = \frac{x}{A^{1+K\delta+o(1)}}.\end{aligned}$$

Since we can take δ arbitrarily small, the proposition follows. □

Acknowledgments. The author would like to acknowledge numerous helpful suggestions made by Kevin Ford, Carl Pomerance, and the referee.

References

[1] N. L. Bassily, I. Kátai, and M. Wijsmuller, *On the prime power divisors of the iterates of the Euler-ϕ function,* Publ. Math. Debrecen 55 (1999), 17–32.

[2] P. Erdős, *On perfect and multiply perfect numbers,* Ann. Mat. Pura Appl. (4) 42 (1956), 253–258.

[3] P. Erdős, F. Luca, and C. Pomerance, *On the proportion of numbers coprime to a given integer,* Anatomy of integers, CRM Proc. Lecture Notes, 46, pp. 47–64, Amer. Math. Soc., Providence, RI, 2008.

[4] P. Erdős and J.-L. Nicolas, *Sur la fonction "nombre de facteurs premiers de n",* Séminaire Delange-Pisot-Poitou, 20e année: 1978/1979. Théorie des nombres, fasc. 2 (French), Exp. no. 32, Secrétariat Math., Paris, 1980.

[5] H. Halberstam and H.-E. Richert, *Sieve methods,* London Math. Soc. Monogr. (N.S.), 4, Academic Press, London, 1974.

[6] G. H. Hardy and E. M. Wright, *An introduction to the theory of numbers,* 5th ed., Oxford Univ. Press, New York, 1979.

[7] A. Ivić and G. Tenenbaum, *Local densities over integers free of large prime factors,* Quart. J. Math. Oxford Ser. (2) 37 (1986), 401–417.

[8] I. Kátai, *On the number of prime factors of* $\phi(\phi(n))$, Acta Math. Hungar. 58 (1991), 211–225.

[9] I. Kátai and M. V. Subbarao, *Some further remarks on the ϕ and σ-functions,* Ann. Univ. Sci. Budapest. Sect. Comput. 26 (2006), 51–63.

[10] F. Luca and C. Pomerance, *Irreducible radical extensions and Euler-function chains,* Combinatorial number theory, pp. 351–361, de Gruyter, Berlin, 2007.

[11] L. Lucht, *Über die Hintereinanderschaltung mulitplikativer Funktionen,* J. Reine Angew. Math. 283/284 (1976), 275–281.

[12] M. Naimi, *Les entiers sans facteurs carré $\leq x$ dont leurs facteurs premiers $\leq y$,* Groupe de travail en théorie analytique et élémentaire des nombres, 1986–1987, Publ. Math. Orsay, 88, pp. 69–76, Univ. Paris XI, Orsay, 1988.

[13] C. Pomerance, *On the distribution of amicable numbers,* J. Reine Angew. Math. 293/294 (1977), 217–222.

[14] G. Tenenbaum, *Introduction to analytic and probabilistic number theory,* Cambridge Stud. Adv. Math., 46, Cambridge Univ. Press, Cambridge, 1995 [translated from the second French edition (1995) by C. B. Thomas].

[15] E. Wirsing, *Bemerkung zu der Arbeit über vollkommene Zahlen,* Math. Ann. 137 (1959), 316–318.

Department of Mathematics
University of Illinois at Urbana-Champaign
Urbana, IL 61801

pppollac@illinois.edu