ACTA ARITHMETICA
146.1 (2011)

Long gaps between deficient numbers

by

PAUL POLLACK (Urbana, IL)

1. Introduction. The ancient Greeks called the positive integer n *deficient*, *perfect*, or *abundant*, according to whether $\sigma(n) < 2n$, $\sigma(n) = 2n$, or $\sigma(n) > 2n$, respectively. Here $\sigma(n) := \sum_{d \mid n} d$ is the usual sum of divisors function. Denoting these sets $\mathcal{D}$, $\mathcal{P}$, and $\mathcal{A}$, we have

$$\mathcal{D} = \{1, 2, 3, 4, 5, 7, 8, 9, 10, 11, 13, 14, 15, 16, 17, \dots\},$$
$$\mathcal{P} = \{6, 28, 496, 8128, 33\,550\,336, 8\,589\,869\,056, 137\,438\,691\,328, \dots\},$$
$$\mathcal{A} = \{12, 18, 20, 24, 30, 36, 40, 42, 48, 54, 56, 60, 66, 70, 72, \dots\}.$$

From the analytic standpoint, it is natural to ask what proportion of the natural numbers fall into each of these three classes. From the limited evidence presented above, we might conjecture that the perfect numbers have density zero, while the abundant and deficient numbers each make up a positive proportion of the integers.

Actually our innocent question requires some care to make precise, as it is not clear a priori that these proportions are well-defined. That this is the case follows from a result discovered independently by each of Behrend, Chowla [1], and Davenport [2]:

THEOREM A. *For each real number u, define*

$$D(u) := \lim_{x \to \infty} \frac{\#\{n \le x : \sigma(n)/n \le u\}}{x}.$$

Then $D(u)$ exists for all u. Moreover, $D(u)$ is a continuous function of u and is strictly increasing for $u \ge 1$. Finally, $D(1) = 0$ and $\lim_{u \to \infty} D(u) = 1$.

The continuity of $D(u)$ implies immediately that the perfect numbers make up a set of asymptotic density zero. We then deduce that the deficient numbers make up a set of asymptotic density $D(2)$ and that the abundant

2010 *Mathematics Subject Classification*: Primary 11A25; Secondary 11N60.
Key words and phrases: aliquot parts, deficient numbers, abundant numbers, additive arithmetic functions.

DOI: 10.4064/aa146-1-3

numbers make up a set of asymptotic density $1 - D(2)$. Deléglise [4] has shown that

$$0.2474 < 1 - D(2) < 0.2480.$$

Thus just under 1 in 4 natural numbers are abundant.

These results adequately describe the global distribution of the deficient and abundant numbers, but it is reasonable to ask also about the local distribution. For abundant numbers, this is straightforward: For example, every proper multiple of 6 is abundant, so the maximal gap between consecutive abundant numbers is at most 6. Moreover, by an elementary averaging argument (cf. [14]), each gap between 1 and 6 can be shown to occur a positive proportion of the time (i.e., with positive lower density). It follows from a theorem of Erdős and Schinzel [8, Theorem 3] that each of these proportions tends to a well-defined limit (i.e., the densities in question exist). For deficient numbers the situation is less simple to describe. For $x \geq 2$, define $G(x)$ as the largest gap $n' - n$ between consecutive deficient numbers $n < n' \leq x$. In 1935, Erdős [5] showed the existence of positive constants c_1 and c_2 with

$$c_1 \log\log\log x \leq G(x) \leq c_2 \log\log\log x \tag{1}$$

for large x. The results of [5], sometimes in slightly weaker form, have been rediscovered multiple times; see, e.g., [9], [10], [12], [3]. Other results on the local distribution of $\sigma(n)/n$ are considered in [7].

Our primary objective is to fill the gap implicit in (1) by proving an asymptotic formula for $G(x)$ as $x \to \infty$.

Theorem 1. *We have $G(x)/\log\log\log x \to 1/C$ as $x \to \infty$, where*

$$C := \int_1^2 \frac{D(u)}{u}\, du. \tag{2}$$

Remark. M. Kobayashi shows [11] that

$$0.28209 \leq C \leq 0.28724, \quad \text{so that} \quad 3.481 \leq \frac{1}{C} \leq 3.545.$$

Erdős originally phrased his theorem in terms of long runs of abundant numbers. We now turn our attention to the question of how often such long runs occur. If $x \geq 2$ is real and A is a positive integer, we let $N(x, A)$ denote the number of $n \leq x$ for which $n + 1, \ldots, n + A$ are all nondeficient (perfect or abundant). The result of Erdős and Schinzel alluded to above implies that for any fixed A, the ratio $N(x, A)/x$ tends to a limit as $x \to \infty$. Our second result shows that this ratio decays triply exponentially with A.

Theorem 2. *With C as defined in (2), we have*

$$N(x, A) \ll \frac{x}{\exp\exp\exp((C + o(1))A)}.$$

Here "$o(1)$" indicates a term that tends to zero as $A \to \infty$ (uniformly in $x \geq 2$), and the constant implied by "$\ll$" is absolute.

Theorem 2 implies half of Theorem 1, namely that

$$\limsup G(x)/\log\log\log x \leq 1/C.$$

The plan for the rest of the paper is as follows: In §2, we prove some lemmas which are useful in the proofs of both theorems. In §3 we prove Theorem 2, and in §4 we complete the proof of Theorem 1 by proving that

$$\liminf G(x)/\log\log\log x \geq 1/C.$$

2. Preparation. We begin by recording the following lemma, which is a special case of [13, Satz I]. Let $D^*(t)$ denote the density of n with $n/\sigma(n) \leq t$, so that $D^*(t) = 1 - D(1/t)$ in the notation of Theorem A.

Lemma 1. *Suppose that f is a function of bounded variation on $[0,1]$. Then, as $A \to \infty$,*

$$\frac{1}{A}\sum_{n\leq A} f\left(\frac{n}{\sigma(n)}\right) \to \int_0^1 f(t)\,dD^*(t).$$

Given a natural number B, define F_B as the arithmetic function which returns the B-*smooth part* of its argument, so that $F_B(n) := \prod_{p\leq B} p^{v_p(n)}$. (Here we write $v_p(n)$ for the exponent with the property that $p^{v_p(n)} \parallel n$.) Put

$$H(n) := \log\frac{\sigma(n)}{n} \quad\text{and}\quad H_B(n) := \log\frac{\sigma(F_B(n))}{F_B(n)},$$

so that $H_B = H \circ F_B$. Note that both H and H_B are additive functions and that $H \geq H_B$ pointwise. H_B has an important near-periodicity property, which we state precisely in the following lemma:

Lemma 2. *Suppose that m and m' are natural numbers with $m \equiv m' \pmod{M}$, where $M := (\prod_{p\leq B} p)^B$. Then $H_B(m) - H_B(m') \to 0$ as $B \to \infty$, uniformly in m and m'.*

Proof. Let $p \leq B$ be prime. Since $v_p(M) = B$ and $M \mid m - m'$, either $v_p(m) = v_p(m')$ or both $v_p(m) \geq B$ and $v_p(m') \geq B$. Consequently,

$$\begin{aligned} H_B(m) - H_B(m') &= \sum_{p\leq B}\log\frac{\sigma(p^{v_p(m)})}{p^{v_p(m)}} - \sum_{p\leq B}\log\frac{\sigma(p^{v_p(m')})}{p^{v_p(m')}} \\ &= \sum_{\substack{p\leq B\\ v_p(m)\geq B,\, v_p(m')\geq B}} \log\left(\frac{p^{v_p(m)+1}-1}{p^{v_p(m)}}\,\frac{p^{v_p(m')}}{p^{v_p(m')+1}-1}\right) \ll \sum_p \frac{1}{p^{B+1}}, \end{aligned}$$

which tends to zero as $B \to \infty$ (by dominated convergence). ∎

Lemma 3. *When $B \to \infty$, we have*

$$\frac{1}{M}\sum_{m=1}^{M}\max\{\log 2 - H_B(m), 0\} \to \int_1^2 \frac{D(u)}{u}\,du,$$

where $M = (\prod_{p\le B} p)^B$, as above.

Proof. By Lemma 1 with $f(x) := \max\{\log 2 - \log(1/x), 0\}$ (interpreted so that $f(0) = 0$), we have

$$\frac{1}{M}\sum_{m=1}^{M}\max\{\log 2 - H(m), 0\} \to \int_{1/2}^{1} (\log 2 + \log u)\,dD^*(u) \tag{3}$$

$$= \log 2 - \int_{1/2}^{1} \frac{D^*(u)}{u}\,du = \int_{1/2}^{1} \frac{1 - D^*(u)}{u}\,du = \int_1^2 \frac{D(u)}{u}\,du$$

as B, and hence M, tends to infinity. Now notice that

$$0 \le \frac{1}{M}\sum_{m\le M}(\max\{\log 2 - H_B(m), 0\} - \max\{\log 2 - H(m), 0\})$$

$$\le \frac{1}{M}\sum_{m\le M}(H(m) - H_B(m)) = \frac{1}{M}\sum_{m\le M}\sum_{\substack{p\mid m\\ p>B}}\log\frac{\sigma(p^{v_p(m)})}{p^{v_p(m)}}.$$

Since

$$\log\frac{\sigma(p^{v_p(m)})}{p^{v_p(m)}} = \log\left(1 + \frac{1}{p} + \cdots + \frac{1}{p^{v_p(m)}}\right) \tag{4}$$

$$< \log\left(1 + \frac{1}{p-1}\right) < \frac{1}{p-1},$$

the above quantity is bounded above by

$$\frac{1}{M}\sum_{m\le M}\sum_{\substack{p\mid m\\ p>B}}\frac{1}{p-1} = \frac{1}{M}\sum_{p>B}\frac{1}{p-1}\sum_{\substack{m\le M\\ p\mid m}} 1 \le \sum_{p>B}\frac{1}{p(p-1)} \le \frac{1}{B}.$$

Since $B \to \infty$, the result follows from (3). ∎

Combining Lemmas 2 and 3 yields the key estimate for the proofs of both Theorems 1 and 2.

Lemma 4. *For each $A > 1$, put $B := \lfloor(\log A)^{1/3}\rfloor$. As $A \to \infty$, we have*

$$\frac{1}{A}\sum_{m=n+1}^{n+A}\max\{\log 2 - H_B(m), 0\} \to \int_1^2 \frac{D(u)}{u}\,du,$$

uniformly for nonnegative integers n.

Proof. Put $M = (\prod_{p \le B} p)^B$, and notice that by the prime number theorem (or a more elementary estimate), $M \le \exp(O((\log A)^{2/3}))$. In particular, $M = o(A)$. We split the sum in the lemma into a sum over blocks of the form $Mk + 1, Mk + 2, \dots, M(k + 1)$, excluding $O(M)$ terms at the beginning and end of the original range of summation. By Lemma 3 and the near-periodicity established in Lemma 2, we find that

$$\sum_{m=Mk+1}^{M(k+1)} \max\{\log 2 - H_B(m), 0\} = (C + o(1))M,$$

where $C = \int_1^2 D(u)u^{-1}\,du$. Since there are $A/M + O(1)$ blocks, the total contribution from all the blocks is $(C + o(1))A$. Finally, notice that the contribution from the $O(M)$ excluded initial and final terms is $O(M) = o(A)$, since each summand is bounded by $\log 2$. ■

We conclude this section with an estimate for certain reciprocal sums, which is implicit in much of Erdős's work:

Lemma 5. *Let $\mathcal{Q}$ be a finite set of primes. Then for each integer $k \ge 0$,*

$$\sum_{\substack{n \text{ squarefree} \\ p \mid n \Rightarrow p \in \mathcal{Q} \\ \omega(n) = k}} \frac{1}{n} \le \frac{1}{k!}\Big(\sum_{p \in \mathcal{Q}} \frac{1}{p}\Big)^k.$$

Proof. If $p_1, \dots, p_k$ are k distinct elements of $\mathcal{Q}$, then by the multinomial theorem, the term $(p_1 \cdots p_k)^{-1}$ appears $\frac{k!}{1!1!\cdots 1!} = k!$ times in the expansion of the right-hand side. Thus the right-hand side majorizes the left. ■

3. Proof of Theorem 2. By adjusting the implied constant, we can assume that A is large. We can also assume that $A \le x$, since otherwise $N(x, A) = 0$. To see this last claim, notice that when $A > x > n$, the interval $n + 1, \dots, n + A$ contains not merely a deficient number but in fact a prime number. This follows from *Bertrand's postulate*, which asserts the existence of a prime in every interval of the form $(n, 2n]$.

So suppose that n is counted by $N(x, A)$ where A is large but $A \le x$. Put $B = \lfloor (\log A)^{1/3} \rfloor$. Then by (4) and Lemma 4,

$$\sum_{m=n+1}^{n+A} \sum_{\substack{p \mid m \\ p > B}} \frac{1}{p-1} \ge \sum_{m=n+1}^{n+A} \sum_{\substack{p \mid m \\ p > B}} \log \frac{\sigma(p^{v_p(m)})}{p^{v_p(m)}} \tag{5}$$

$$= \sum_{m=n+1}^{n+A} (H(m) - H_B(m)) \ge \sum_{m=n+1}^{n+A} \max\{\log 2 - H_B(m), 0\} \ge (C + o(1))A,$$

where C is given by (2). To proceed we need a lower bound on $\sum_{p\mid N}\frac{1}{p-1}$, where

$$N := (n+1)(n+2)\cdots(n+A). \tag{6}$$

To obtain our bound, we remove the overlap from (5), corresponding to primes p which divide more than one of $n+1,\dots,n+A$. Clearly any such p satisfies $p < A$, and so the contribution of such primes to the sum in (5) is bounded by

$$\sum_{B<p<A}\frac{1}{p-1}\sum_{\substack{n+1\le m\le n+A\\ p\mid m}} 1 \le 2A\sum_{p>B}\frac{1}{p(p-1)} \le \frac{2A}{B},$$

which is $o(A)$ as $A\to\infty$. It follows that we can choose a function $r(A)$, depending only on A, with $r(A)=o(1)$ as $A\to\infty$ and

$$\sum_{p\mid N}\frac{1}{p-1} \ge (C-r(A))A.$$

Let $Z = Z(A)$ be the smallest positive integer for which

$$\sum_{p\le Z}\frac{1}{p-1} \ge (C-r(A))A, \quad \text{so} \quad Z = \exp(\exp((C+o(1))A)) \quad \text{as } A\to\infty.$$

(Here we use the well-known estimate $\sum_{p\le Z}(1/p) = \log\log Z + O(1)$.) We now split the n counted by $N(x,A)$ into two classes:

(i) In the first class, we consider those n for which N, as defined in (6), has at least $Z/(2\log Z)$ distinct prime divisors not exceeding $4Z$.
(ii) In the second class we put all the remaining values of n.

If n belongs to the first class, then for some $1\le i\le A$, the number $n+i$ has (at least) $(2A)^{-1}Z/\log Z$ prime divisors not exceeding $4Z$. Since $n+i\le 2x$, the number of possible n that arise this way is at most

$$A\cdot\frac{2x}{k!}\Bigg(\sum_{p\le 4Z}\frac{1}{p}\Bigg)^k \quad \text{where} \quad k := \left\lceil \frac{1}{2A}\frac{Z}{\log Z}\right\rceil,$$

by Lemma 5. A short calculation shows that for large A, this bound is

$$\ll xA\frac{\exp\big(O\big(Z\frac{\log\log\log Z}{\log Z\log\log Z}\big)\big)}{\exp(Z/(4A))} \ll \frac{x}{\exp\exp\exp((C+o(1))A)}.$$

(To verify this, it is helpful to keep in mind that $A \asymp \log\log Z$ for large A.)

It remains to show that we have a similar estimate for the n belonging to the second class. Note that for large Z, the first $Z/(2\log Z)$ primes all belong to the interval $[1, 2Z/3]$. Consequently, for n in the second class,

$$\sum_{\substack{p \mid N \\ p \le 4Z}} \frac{1}{p-1} \le \sum_{p \le 2Z/3} \frac{1}{p-1},$$

once A is sufficiently large. Hence,

$$\begin{aligned} \sum_{\substack{p \mid N \\ p > 4Z}} \frac{1}{p-1} &\ge \sum_{p \mid N} \frac{1}{p-1} - \sum_{p \le 2Z/3} \frac{1}{p-1} \\ &\ge (C - r(A))A - \sum_{p \le 2Z/3} \frac{1}{p-1} \ge \sum_{2Z/3 < p < Z} \frac{1}{p-1}, \end{aligned} \tag{7}$$

where for the last inequality we use the minimality of Z. Now

$$\sum_{2Z/3 < p < Z} \frac{1}{p-1} \ge \frac{1}{Z}(\pi(Z-1) - \pi(2Z/3)) > \frac{1}{4 \log Z} \tag{8}$$

when A, and hence Z, is large. It follows from (7) and (8) that there is some $j \ge 1$ for which

$$\sum_{\substack{p \mid N \\ 4^j Z < p \le 4^{j+1} Z}} \frac{1}{p-1} \ge \frac{1}{2^j} \frac{1}{4 \log Z}.$$

For this j, the number N is divisible by at least $2^j Z/(4 \log Z)$ primes from the interval $(4^j Z, 4^{j+1} Z]$, and so one of the numbers $n+1, \ldots, n+A$ is divisible by at least

$$W_j := \left\lceil \frac{2^j Z}{4A \log Z} \right\rceil$$

of these primes. By another application of Lemma 5, we deduce that the number of such n is at most

$$A \frac{2x}{W_j!} \Bigg(\sum_{4^j Z < p \le 4^{j+1} Z} \frac{1}{p} \Bigg)^{W_j}.$$

Now we sum this expression over j, noting that for large A the inner sum here is bounded by $1/2$. This gives (for large A) an upper bound on the number of n in the second class which is

$$\ll Ax \sum_{j=1}^{\infty} \frac{1}{2^{W_j} W_j!}.$$

For large A, the sum here is dominated by its first term, and we obtain a final bound of

$$\ll \frac{Ax}{2^{W_1} W_1!} = \frac{Ax}{\exp((1+o(1))W_1 \log W_1)} \leq \frac{Ax}{\exp(Z/(4A))}$$
$$\ll \frac{X}{\exp\exp\exp((C+o(1))A)},$$

as desired.

REMARK. This proof uses a method introduced by Erdős [6] and applied by him to estimate $D(u)$ as $u \to \infty$. In this connection, see the recent paper [15].

4. Proof of Theorem 1. Fix $\epsilon > 0$, which we assume to be small. For all large x, we will construct a positive integer $n \leq x/2$ with $n+1, n+2, \ldots, n+A$ all nondeficient, where

$$A := \lceil (C+8\epsilon)^{-1} \log\log\log x \rceil. \tag{9}$$

If n' denotes the first deficient number after n, then (for example, by Bertrand's postulate) $n' \leq 2n \leq x$, and $n'-n > (n+A)-n = A$. Thus $G(x) > A$. Since $\epsilon > 0$ is arbitrary, this implies the lower bound implicit in Theorem 1. (Recall that the upper bound follows from Theorem 2.)

With $B := \lfloor (\log A)^{1/3} \rfloor$, we let $p_0 < p_1 < \cdots$ be the sequence of consecutive primes exceeding B. Put $i_0 := 0$. If $i_0, i_1, \ldots, i_{l-1}$ have been defined, choose i_l as small as possible so that

$$H(P_l) \geq \epsilon + \max\{\log 2 - H_B(l), 0\} \quad \text{where} \quad P_l := \prod_{i_{l-1} \leq p_j < i_l} p_j.$$

Suppose now that n is chosen so that

$$M := \Big(\prod_{p \leq B} p\Big)^B \,|\, n \quad \text{and} \quad P_l \,|\, n+l \quad \text{for all } 1 \leq l \leq A.$$

Such a choice is possible by the Chinese remainder theorem; indeed, the n which satisfy these conditions make up a nonzero residue class modulo $M \prod_{l=1}^{A} P_l$. For any such n, and any $1 \leq l \leq A$, we have

$$H(n+l) \geq H(P_l) + H_B(n+l) \geq H(P_l) + H_B(l) - \epsilon \geq \log 2,$$

once x is sufficiently large. (Here we use the near-periodicity property established in Lemma 2 to obtain the last inequality.) Thus all of $n+1, \ldots, n+A$ are nondeficient.

To show that such a choice is possible with $n \leq x/2$, it is enough to show that

$$M \prod_{l=1}^{A} P_l \leq x/2,$$

once x is sufficiently large. Write $\prod_{l=1}^{A} P_l$ in the form $\prod_{B<p\leq Z} p$, where Z

is chosen as small as possible. Then for large A,

$$\sum_{B<p\le Z}\log\frac{\sigma(p)}{p}=\sum_{l=1}^{A}H(P_l)\le\sum_{l=1}^{A}(\max\{\log 2-H_B(l),0\}+\epsilon+1/B)$$
$$\le AC+2A\epsilon+A/B=A(C+3\epsilon).$$

(Here the last inequality of the first line follows from our choosing i_l minimally at each stage.) Since $\log(\sigma(p)/p)=1/p+O(1/p^2)$, it follows that

$$\sum_{p\le Z}\frac{1}{p}\le A(C+3\epsilon)+O(1)+\log\log B\le A(C+4\epsilon),$$

and hence

$$Z\le\exp\exp(A(C+5\epsilon)).$$

Consequently,

$$M\prod_{1\le l\le A}P_l\le\exp(O((\log A)^{2/3}))\prod_{p\le Z}p$$
$$\le\exp(O((\log A)^{2/3}))\exp\exp\exp(A(C+6\epsilon))$$
$$\le\exp\exp\exp(A(C+7\epsilon))\le x/2$$

by our definition (9) of A.

Remark. The above argument shows that for each $\epsilon>0$ and all large x, we have

$$N(x,A)\ge\left\lfloor\frac{x}{M\prod_{1\le l\le A}P_l}\right\rfloor\ge x/\exp\exp\exp(A(C+7\epsilon)) \tag{10}$$

as $x\to\infty$, if A is defined by (9). In fact, our proof shows that if A tends to infinity with x and if A is bounded above by the expression on the right of (9), then the inequality (10) holds. This can be viewed as a lower-bound counterpart of Theorem 2.

Acknowledgements. This material is based upon work supported by the National Science Foundation under agreement No. DMS-0635607. I would like to thank Carl Pomerance for suggesting the problem of determining an asymptotic formula for $G(x)$. I am also grateful to Mitsuo Kobayashi for estimating the value of C, supplying the reference for Lemma 1, and for giving nice talks that influenced the writing of the introduction. Finally, the referee deserves thanks for several helpful suggestions.

References

[1] S. Chowla, *On abundant numbers*, J. Indian Math. Soc. 1 (1934), 41–44.
[2] H. Davenport, *Über numeri abundantes*, Sitzungsber. Akad. Berlin 1933, 830–837.

[3] J.-M. De Koninck and I. Kátai, *On the frequency of k-deficient numbers*, Publ. Math. Debrecen 61 (2002), 595–602.
[4] M. Deléglise, *Bounds for the density of abundant integers*, Experiment. Math. 7 (1998), 137–143.
[5] P. Erdős, *Note on consecutive abundant numbers*, J. London Math. Soc. 10 (1935), 128–131.
[6] —, *Some remarks about additive and multiplicative functions*, Bull. Amer. Math. Soc. 52 (1946), 527–537.
[7] —, *Some remarks on Euler's φ function*, Acta Arith. 4 (1958), 10–19.
[8] P. Erdős and A. Schinzel, *Distributions of the values of some arithmetical functions*, ibid. 6 (1960/1961), 473–485.
[9] J. Galambos, *On a conjecture of Kátai concerning weakly composite numbers*, Proc. Amer. Math. Soc. 96 (1986), 215–216.
[10] J. Galambos and I. Kátai, *The gaps in a particular sequence of integers of positive density*, J. London Math. Soc. (2) 36 (1987), 429–435.
[11] M. Kobayashi, private communication.
[12] J. Sándor, *On a method of Galambos and Kátai concerning the frequency of deficient numbers*, Publ. Math. Debrecen 39 (1991), 155–157.
[13] I. Schoenberg, *Über die asymptotische Verteilung reeler Zahlen mod 1*, Math. Z. 28 (1928), 171–199.
[14] C. R. Wall, *Problem 6356*, Amer. Math. Monthly 88 (1981), 623 (solution by L. L. Foster, ibid. 90 (1983), 215–216).
[15] A. Weingartner, *The distribution functions of $\sigma(n)/n$ and $n/\varphi(n)$*, Proc. Amer. Math. Soc. 135 (2007), 2677–2681.

Paul Pollack
Department of Mathematics
University of Illinois
1409 West Green Street
Urbana, IL 61801, U.S.A.
E-mail: pppollac@illinois.edu
http://www.math.uiuc.edu/~pppollac/

Received on 12.11.2009
and in revised form on 21.1.2010 (6215)