REMARKS ON AN ARITHMETIC DERIVATIVE

E. J. Barbeau

(received March 21, 1961)

1. Introduction. Let $D(n)$ denote a function of an integral variable $n \geq 0$ such that¹

1. $D(1) = D(0) = 0$

2. $D(p) = 1$ for every prime $p$

3. $D(n_1 n_2) = n_1 D(n_2) + n_2 D(n_1)$ for every pair of non-negative integers $n_1, n_2$.

The property (3) is analogous to the product rule for derivatives, and its extension to $k$ terms

4. $D(n) = \sum_{i=1}^{k} n_i^{-1} D(n_i)$ for $n = n_1 n_2 \cdots n_k$

is immediate. The above properties are consistent and determine $D(n)$ uniquely for all non-negative integers $n$. In fact, if $n = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_r^{\alpha_r}$, we have, on using (4),

5. $D(n) = \sum_{i=1}^{r} \alpha_i p_i^{-1}$

so that, once the prime factor decomposition of $n$ is known, the first derivative $D(n)$ is given explicitly. However, the "higher" derivatives, defined successively by

$D^0(n) = n, \ D^1(n) = D(n), \ D^2(n) = D[D(n)], \ldots, \ D^k(n) = D[D^{k-1}(n)]$

¹I have not been able to trace explicit references to previous work on $D(n)$. However, it appeared in a question on the Putnam Prize competition (1950); see American Mathematical Monthly 57 (1950), p. 469. I am indebted to Dr. J. H. H. Chalk for suggesting a note on this topic and for assistance during its preparation.


117

https://doi.org/10.4153/CMB-1961-013-0 Published online by Cambridge University Press
present an unsolved problem. For fixed \( n \), the function \( D_k(n) \) of \( k \) exhibits irregular behaviour as \( k \) increases. For example, using (3) with \( n = p^Pn_1 \), where \( p \) is a prime, we obtain

\[
D(n) = p^P[n_1 + D(n_1)] \geq n
\]
equality holding if and only if \( n_1 = 1 \). Hence, for integers \( n \) possessing a proper divisor of the form \( p^P \), \( \lim D_k(n) = \infty \), and if \( n = p^P \), \( D_k(n) = n \) for all \( k \). On the other hand, \( D_k(p) = 0 \) for all \( k > 1 \) and all primes \( p \). Numerical considerations suggest the following.

CONJECTURE. For each \( n > 1 \), there exists a constant \( k_0 = k_0(n) > 1 \) such that, for all \( k > k_0 \),
either

1) \( D_k(n) = 0 \)

or

2) \( D_k(n) \neq 0 \),

and there exists a prime \( p \) such that \( D_k(n) \equiv 0 \pmod{p} \).

2. Some remarks about \( D(n) \). Although the function \( D(n) \) behaves erratically, it is easy to obtain exact upper and lower bounds, depending on \( n \), for its values. We suppose that \( n = q_1 q_2 \ldots q_v \) has prime factors \( q_i \) which are not necessarily distinct.

(a) \( D(n) < \frac{n \log n}{2 \log 2} \) for all \( n \), equality occurring if and only if \( n \) is a power of 2. In fact, \( n \) satisfies \( 2^k \leq n < 2^{k+1} \) for some \( k \). Clearly, \( v \leq k \) and

\[
D(n) = n \sum_{i=1}^{v} q_i^\frac{1}{q_i^n} \leq n \sum_{i=1}^{v} \frac{1}{q_i} \leq \frac{n k}{2} \leq \frac{n \log n}{2 \log 2}.
\]

If \( n = 2^k \), \( D(n) = k2^{k-1} = \frac{2^k \log 2}{2 \log 2} \). If \( n \neq 2^k \), then some \( q_i \neq 2 \) and strict inequality holds in the above.

(b) \( D(n) \geq \nu n - \frac{1}{\nu} \), equality holding if, and only if, all the factors \( q_i \) are equal. For, by (5) and the inequality of the arithmetic and geometric means,
\[ D(n) = n \sum_{i=1}^{\nu} \frac{1}{q_i} \geq n^{\nu} \frac{1}{(q_1 q_2 \ldots q_{\nu})^{1/\nu}} = \nu n^{1 - 1/\nu}. \]

Hence, if \( n \) is not a prime or unity, \( D(n) \geq 2\sqrt{n} \), with equality if and only if \( n = p^2 \) where \( p \) is a prime.

In addition, we can relate the value of \( D(n) \) to \( n \) in the following ways.

(c) Let \( n = p_1^{\alpha_1} \ldots p_r^{\alpha_r} \), where \( p_1, \ldots, p_r \) are distinct primes. Then \( D(n) \equiv 0 \) (mod \( n \)) if, and only if, 
\[ \alpha_1 \equiv 0 \pmod{p_1}, \ldots, \alpha_r \equiv 0 \pmod{p_r}. \]
In particular, \( D(n) = n \) if and only if, \( n = p^p \). The sufficiency of the conditions is obvious.

Their necessity is seen by noting that, if \( n = p^{\alpha} n' \), where \( (p, n') = 1 \), then \( D(n) = n'^{\alpha p^{\alpha-1}} + p^{\alpha} D(n') \equiv 0 \pmod{n} \) implies 
\( n'^{\alpha p^{\alpha-1}} = 0 \pmod{p^{\alpha}} \) and, hence, \( \alpha \equiv 0 \pmod{p} \), since \( (n', p) = 1 \).

(d) If \( D(n) > n \), then \( D(kn) = kD(n) + nD(k) > kn \) for all \( k > 1 \).

3. The average order of \( D(n) \). Let
\[ S(n) = \sum_{r=1}^{n} D(r), \quad T(n) = \sum_{r=1}^{n} K(r) \]
where \( K(n) = n^{-1} D(n) \). Since \( K(n) \) is totally additive, i.e.
\( K(n_1 n_2) = K(n_1) + K(n_2) \) for all integer pairs \( n_1, n_2 \), it is easier to estimate \( T(n) \) first, and then use partial summation to deduce the average order of \( D(n) \). Let
\[ j(n, p) = \sum_{t=1}^{\infty} \left[ \frac{n}{p^t} \right] \], \( a(n) = \left[ \frac{\log n}{\log 2} \right] \);
then \( j(n, p) \) denotes \([1; p. 342]\) the exponent of the highest power of \( p \) dividing \( n! \) and \( a(n) \) denotes the exponent of the highest power of \( 2 \) \( \leq n \). Observe that
\[ T(n) = K(n!) = \sum_{p \leq n} \frac{1}{p} j(n, p) \]
\[ = \sum_{p \leq n} \frac{1}{p} (\sum_{t=1}^{\infty} \left[ \frac{n}{p^t} \right]) \]

119
\[
\sum_{p \leq n} \frac{1}{p} \left( \sum_{t=1}^{\alpha(n)} \left[ \frac{n}{p^t} \right] \right)
\]

\[
= \sum_{p \leq n} \frac{1}{p} \left\{ \sum_{t=1}^{\alpha(n)} \frac{n}{p^t} + O(\log n) \right\}
\]

\[
= \sum_{p \leq n} \left\{ \sum_{t=2}^{\infty} \frac{n}{p^t} - \sum_{t=1}^{\alpha(n)+1} \frac{n}{p^t} \right\} + O((\log n) \sum_{p \leq n} \frac{1}{p})
\]

\[
= n \sum_{p \leq n} \frac{1}{p(p-1)} - \sum_{p \leq n} \frac{n}{p^2} + O((\log n) \sum_{p \leq n} \frac{1}{p})
\]

\[
= n \sum_{p=2}^{\infty} \frac{1}{p(p-1)} - \sum_{p > n} \frac{n}{p(p-1)} - \sum_{p \leq n} \frac{n}{p^2}
\]

\[
+ O((\log n) \sum_{p \leq n} \frac{1}{p})
\]

\[
= T_o n + O((\log n)(\log \log n))
\]

where \( T_o = \sum_{p=2}^{\infty} \frac{1}{p(p-1)} = 0.749 \ldots \)

since

\[
\sum_{p > n} \frac{n}{p(p-1)} < n \sum_{k > n} \frac{1}{k(k-1)} \leq 1,
\]

\[
\log n < \frac{\log n}{p} < \frac{\log n}{\log 2} > \frac{1}{p} > \log 2 > n,
\]

\[
\sum_{p \leq n} \left\{ \frac{1}{p-1} - \frac{1}{p} \right\} \leq 1,
\]

\[
\sum_{p \leq n} \frac{1}{p} = O(\log \log n).
\]

[1; p. 351]

For \( S(n) \), we have

\[
S(n) = \sum_{r=1}^{n} rK(r) = T(n) + \sum_{r=1}^{n-1} \{ T(n) - T(r) \}
\]

\[
= nT(n) - \sum_{r=1}^{n-1} T(r)
\]

\[
= n\{ T_o n + O(n^\delta) \} - T_o \sum_{r=1}^{n-1} r + O(n^{1+\delta})
\]

\[
= T_o n^2 - T_o \frac{n(n-1)}{2} + O(n^{1+\delta})
\]

120
\[ \frac{1}{2} T_0 n^2 + O(n^{1+\delta}) \]

where \( \frac{1}{2} T_0 = 0.374 \ldots \), for each fixed \( \delta > 0 \).

4. The congruence \( D(n) \equiv 0 \pmod{4} \). A key problem is to find a characterization of those numbers for which

\[ \lim_{k \to \infty} D^k(n) = \infty. \]

This limit is known for numbers \( n \) of the form \( p, p^k, kp^k \) where \( p \) is any prime. Further investigation is hampered by the absence of explicit formulae for the higher derivatives. If there were some way of dealing with \( D(m + n) \) for any integers \( m \) and \( n \), then \( D^2(n) \) could be determined from \( D(n) = \sum_{i=1}^{k} F_i \), where \( n = \prod_{i=1}^{k} f_i \), \( F_i = n/f_i \), \( f \) prime.

However, it is known only that, if \( D(m + n) = D(m) + D(n) \), then \( D(km + kn) = D(km) + D(kn) \) for every integer \( k \); in particular, \( D(h) + D(2h) = D(3h) \).

Another approach to the problem is to find a characterization of those numbers, excluding \( p, p^k, kp^k \) for which \( p^k \mid D^k(n) \) for some positive integer \( k \) and some prime \( p \). According to our conjecture, this would be sufficient to characterize those numbers for which \( D^k(n) \to \infty \) as \( k \to \infty \), provided \( D^k(n) \neq 0 \) for all \( k \). We deal with the special case \( p = 2, k = 1 \).

Let \( n = 2^a p_1 p_2 \ldots p_r q_1 q_2 \ldots q_s \) where \( p_i \equiv 1 \pmod{4}, q_j \equiv -1 \pmod{4} \) are primes, not necessarily distinct. We have the following results:

(i) if \( a = 0 \), then \( D(n) \equiv (-1)^{s(r - s)} \pmod{2^2} \)
(ii) if \( a = 1 \), then \( D(n) \equiv (-1)^{s[1 + 2(r - s)]} \equiv (-1)^{r-1} \pmod{2^2} \)
(iii) if \( a > 1 \), then \( D(n) \equiv 0 \pmod{2^2} \).

In order to prove (i), let \( P = p_1 p_2 \ldots p_r \equiv (+1) \pmod{4} \)
\[ Q = q_1 q_2 \ldots q_s \equiv (-1)^s \pmod{4} \]
\[ P_i = \frac{P}{p_i} \equiv 1 \pmod{4} \]
\[ Q_i = \frac{Q}{q_i} \equiv (-1)^{s-1} \pmod{4}. \]

The approximation \( 0.374 \ldots n^2 \) for \( S(n) \) is good, even for small values of \( n \). For example, \( S(10) = 38 \equiv (0.374 \ldots)(100) \).
Then
\[ D(n) = D(PQ) = \sum_{i=1}^{r} P_i Q + \sum_{i=1}^{s} P_i Q_i \equiv r(-1)^s + s(-1)^{s-1} \]
\[ \equiv (-1)^s (r - s) \pmod{4}. \]

In case (ii),
\[ D(2PQ) = PQD(2) + 2D(PQ) \]
\[ \equiv (-1)^s + 2(-1)^s (r - s) \]
\[ \equiv (-1)^s [1 + 2(r - s)] \pmod{4}. \]

Result (iii) follows from the fact that \( 4 \mid n \). We conclude that
\[ D(n) \equiv 0 \pmod{4} \text{ if and only if} \]
(a) \( a = 0, \ r \equiv s \pmod{4} \)
(b) \( a > 1 \).

The numbers in (a) have a density of \( \frac{1}{8} \) in the integers; those in
(b) have a density of \( \frac{1}{4} \). Hence, those integers \( n \) satisfying
\[ \lim_{k \to \infty} D^k(n) = \infty \] (which include the numbers of (a) and (b))
have a density exceeding \( \frac{3}{8} \). What this density is remains an
open question.

**REFERENCE**


University of Toronto