NUMBERS WITH INTEGER COMPLEXITY CLOSE TO THE LOWER BOUND

HARRY ALTMAN AND JOSHUA ZELINSKY

Abstract. Define $\|n\|$ to be the *complexity* of n, the smallest number of ones needed to write $\|n\|$ using an arbitrary combination of addition and multiplication. John Selfridge showed that $\|n\| \geq 3\log_3 n$ for all n. Define the *defect* of n, denoted $\delta(n)$, to be $\|n\| - 3\log_3 n$; in this paper we present a method for classifying all n with $\delta(n) < r$ for a given r. From this, we derive several consequences. We prove that $\|2^m 3^k\| = 2m + 3k$ for $m \leq 21$ with m and k not both zero, and present a method that can, with more computation, potentially prove the same for larger m. Furthermore, defining $A_r(x)$ to be the number of n with $\delta(n) < r$ and $n \leq x$, we prove that $A_r(x) = \Theta_r((\log x)^{\lfloor r \rfloor + 1})$, allowing us to conclude that the values of $\|n\| - 3\log_3 n$ can be arbitrarily large.

1. Introduction

The complexity of a natural number n is the least number of 1's needed to write it using any combination of addition and multiplication, with the order of the operations specified using parentheses grouped in any legal nesting. For instance, 11 has complexity of 8, since it can be written using 8 ones as $(1+1+1)(1+1+1)+1+1$, but not with any fewer. This notion was introduced by Kurt Mahler and Jan Popken in 1953 [9]. It was later circulated by Richard Guy [5], who includes it as problem F26 in his *Unsolved Problems in Number Theory* [6]. It has been studied by a number of authors, e.g. Rawsthorne [10] and especially Juan Arias de Reyna [4].

Following Arias de Reyna [4] we will denote the complexity of n by $\|n\|$. Notice that for any natural numbers n and m we will have

$$\|1\| = 1, \quad \|n+m\| \leq \|n\| + \|m\|, \quad \|nm\| \leq \|n\| + \|m\|,$$

More specifically, for any $n > 1$, we have

$$\|n\| = \min_{\substack{a,b<n\in\mathbb{N} \\ a+b=n \text{ or } ab=n}} \|a\| + \|b\|.$$

This fact together with $\|1\| = 1$ allows one to compute $\|n\|$ recursively.

Integer complexity is approximately logarithmic; it satisfies the bounds

$$3\log_3 n \leq \|n\| \leq 3\log_2 n, \qquad n > 1.$$

The upper bound can be obtained by writing n in binary and finding a representation using Horner's algorithm. The lower bound follows from results described below. The lower bound is known to be attained infinitely often, namely for all $n = 3^k$. The constant in the upper bound above can be improved further [13], and it is an open problem to determine the true asymptotic order of magnitude

Date: April 16, 2012.

of the upper bound. At present even the possibility that an asymptotic formula $\|n\| \sim 3\log_3 n$ might hold has not been ruled out.

Let $E(k)$ be the largest number writable with k ones, i.e., with complexity at most k. John Selfridge (see [5]) proved that $E(1) = 1$, and the larger values depend on the residue class of k modulo 3, namely for $k = 3j + i \geq 2$,

$$\begin{aligned} E(3j) &= 3^j \\ E(3j+1) &= 4 \cdot 3^{j-1} \\ E(3j+2) &= 2 \cdot 3^j \end{aligned}$$

Observe that $E(k) \geq 3^{k/3}$ in all cases, and that equality holds for cases where 3 divides k. These formulas also show that $E(k) > E(k-1)$, a fact that implies that the integer $E(k)$ requires exactly k ones. This yields the following result:

Theorem 1.1. *For $k \geq 0$, and $a = 0, 1, 2$, with $a + k \geq 1$, one has*

$$\|2^a \cdot 3^k\| = 2a + 3k.$$

Further results are known on the largest possible integers having a given complexity. We can generalize the notion of $E(k)$ with the following definition:

Definition 1.2. Define $E_r(k)$ to be the $(r+1)$-th largest number writable using k ones, i.e. complexity at most k, so long as there are indeed $r + 1$ or more distinct such numbers. Thus $E_r(k)$ is defined only for $k \geq k(r)$. Here $E_0(k) = E(k)$.

Daniel A. Rawsthorne [10] determined a formula for $E_1(k)$, namely:

$$E_1(k) = \frac{8}{9}E(k), \qquad k \geq 8$$

Direct computation establishes that $E_1(k) \leq (8/9)E(k)$ holds for all $2 \leq k \leq 7$ (note that $E_1(1)$ is not defined). From this fact we deduce that, for $0 \leq a \leq 5$ and all $k \geq 0$ with $a + k > 0$,

$$\|2^a \cdot 3^k\| = 2a + 3k.$$

J. Iraids et al. [7] has verified that $\|2^a 3^k\| = 2a + 3k$ for $2 \leq 2^a \cdot 3^k \leq 10^{12}$, so in particular

$$\|2^a\| = 2a, \quad \text{for} \quad 1 \leq a \leq 39.$$

These results together with results given later in this paper lend support to the following conjecture, which was originally formulated as a question in Guy [5].

Conjecture 1.3. *For all $a \geq 0$ and all $k \geq 0$ with $a + k \geq 1$ there holds*

$$||2^a \cdot 3^k|| = 2a + 3k.$$

This conjecture is presented as as a convenient form for summarizing the existing knowledge; there is limited evidence for its truth, and it may well be false. Indeed its truth would imply $\|2^a\| = 2a$, for all a. Selfridge raised this special case in a contrary form, asking the question whether there is some a for which $\|2^a\| < 2a$ (see [5]).

In this paper, we will investigate these questions by looking at numbers n for which the difference $\delta(n) := \|n\| - 3\log_3 n$ is less than a given threshold; these sets we may call numbers with integer complexity close to the lower bound.

1.1. **Main Results.** The fundamental issue making the complexity of an integer a complicated quantity are: (1) It assumes the same value for many integers, because it is logarithmically small; (2) It is hard to determine lower bounds for a given value $\|n\|$, since the dynamic programming tree is exponentially large. The feature (1) implies there can be many tie values in going down the tree, requiring a very large search, to determine any specific complexity value.

We introduce a new invariant to study integer complexity.

Definition 1.4. The *defect* of a natural number n is given by

$$\delta(n) = \|n\| - 3\log_3 n$$

The introduction of the defect simplifies things in that it provides a more discriminating invariant: we show that $\delta(n) \geq 0$ and that it separates integers into quite small equivalence classes. In these equivalence classes powers of 3 play a special role. The following result establishes a conjecture of Arias de Reyna [4, Conjecture 1].

Theorem 1.5. *(1) For a given value δ of the defect, the set $S(\delta) := \{m : \delta(m) = \delta\}$, is a chain $\{n \cdot 3^k : 0 \leq k \leq k(n)\}$ where $k(n)$ may be finite or infinite. The value n is called the leader of the chain.*

(2) The function $\delta(n \cdot 3^k)$ is non-increasing on the sequence $\{n \cdot 3^k : k \geq 0\}$. This sequence has a finite number of leaders culminating in a largest leader $n \cdot 3^L$, having the property that

$$||n \cdot 3^k|| = ||n \cdot 3^L|| + 3(k - L), \text{ for all } k \geq L.$$

The set of integers $n \cdot 3^k$ for $k \geq L$ are termed *stable integers*, because their representation using 1's stabilizes into a predictable form for $k \geq L$. This result is proved in Section 2.1.

The main results of the paper concern classifying integers having small values of the defect. The defect is compatible with the multiplication aspect of the dynamic programming definition of the integer complexity, but it does not fully respect the addition aspect. The main method underlying the results of this paper is given in Theorem 4.4, which provides strong constraints on the dynamic programming recursion for classifying numbers of small defect. It allows construction of sets of integers including all integers of defect below a specified bound r, which may however include some additional integers. The method contains adjustable parameters, and with additional work they sometimes permit exact determination of these sets.

This main method has several applications. First, we use it to explictly classify all integers of defect below the bound $12\delta(2) \approx 1.286$. (Theorem 5.1). This requires pruning the sets found using Theorem 4.4 to determine the sets below $k\delta(2)$ for $1 \leq k \leq 12$.

Using this result we obtain an explicit classification of all integers having defect at most 1, as follows.

Theorem 1.6. *The numbers n satisfying $0 \leq \delta(n) < 1$ are precisely those that can be written in one of the following forms, and have the following complexities:*

(1) 3^k *for* $k \geq 1$, *of complexity* $3k$
(2) $2^a 3^k$ *for* $a \leq 9$, *of complexity* $2a + 3k$ *(for a, k not both zero)*
(3) $5 \cdot 2^a 3^k$ *for* $a \leq 3$, *of complexity* $5 + 2a + 3k$
(4) $7 \cdot 2^a 3^k$ *for* $a \leq 2$, *of complexity* $6 + 2a + 3k$

(5) $19 \cdot 3^k$ *of complexity* $9 + 3k$
(6) $13 \cdot 3^k$ *of complexity* $8 + 3k$
(7) $(3^n + 1)3^k$ *of complexity* $1 + 3n + 3k$ *(for* $n \neq 0$*)*

Furthermore $n = 1$ *is the only number having defect exactly* 1.

This result is established in Section 6.1. Using a slightly more general result, which we present as 5.1, one can obtain a generalization of Rawsthorne's results, consisting of a description of all $E_r(k)$ for every finite $r \geq 0$, valid for all sufficiently large k, depending on r. This answer also depends on the congruence class of $k \pmod 3$. For example, one has $E_2(3k) = \frac{64}{81}E(3k)$, $E_2(3k+1) = \frac{5}{6}E(3k+1)$ and $E_2(3k+2) = \frac{5}{6}E(3k+2)$, all holding for $k \geq 4$. For $E_5(k)$ all three residue classes have different formulas, valid for $k \geq 5$. This generalization will be described elsewhere ([1]).

Secondly, the result can be used to obtain lower bounds on complexity of certain integers, by showing they are excluded from sets containing all integers of complexity at most r. This we use to prove Conjecture 1.3 for $a \leq 21$.

Theorem 1.7. *For all* $0 \leq a \leq 21$ *and any* $k \geq 0$ *having* $a + k \geq 1$*, there holds*

$$\|2^a 3^k\| = 2a + 3k.$$

This result is established in Section 6.2. It is possible to carry out computations establishing the Conjecture 1.3 for larger value of a, as we shall describe in [3].

Thirdly, our main method can be used to estimate the magnitude of numbers below x having a given defect.

Theorem 1.8. *For any* $r > 0$ *the number of elements* $A_r(x)$ *smaller than* x *which have complexity* $\delta(n) < r$ *satisfies an upper bound, valid for all* $x \geq 2$*,*

$$A_r(x) \leq C_r(\log x)^{\lfloor r \rfloor + 1},$$

where $C_r > 0$ *is an effectively computable constant depending on* r*.*

This result is proved in Section 6.3. It implies that the set of possible defect values is unbounded.

1.2. **Discussion.** We first remark on computing $\|n\|$. The recursive definition permits computing $\|n\|$ by dynamic programming, but it requires knowing $\{\|k\| : 1 \leq k \leq n-1\}$, so takes exponential time in the input size of n measured in bits. In particular, a straightforward approach to computing $\|x\|$ requires on the order of n^2 steps. Srinivas and Shankar [11] obtained an improvement on this.

We make some further remarks on Conjecture 1.3. Let's specialize to $k = 0$ and consider an analogous question for prime powers, concerning $\|p^k\|$ as k varies. It is clear that $\|p^k\| \leq k \cdot \|p\|$, since we can concatenate by multiplication k copies of a good representation of p. For which primes p is it true that $\|p^k\| = k\|p\|$ holds for all $k \geq 1$? This is verified for $p = 3$ by $\|3^k\| = 3k$, and the truth of Conjecture 1.3 requires that it hold for $p = 2$, with $\|2^k\| = 2k$. However this question has a negative answer for powers of 5. Here $\|5\| = 5$, one instead gets that $\|5^6\| = \|15625\| = 29 < 6\|5\| = 30$, as

$$\begin{aligned} 15625 \quad = \quad & 1 + (1+1)(1+1)(1+1)(1+1+1)(1+1+1) \cdot \\ & (1 + (1+1)(1+1)(1+1)(1+1+1)(1+1+1)(1+1+1)) \end{aligned}$$

This encodes the identity $5^5 = 1 + 72 \cdot 217$, in which $72 = 2^3 \cdot 3^2$ and $217 = 1 + 2^3 \cdot 3^3$. This counterexample for powers of 5 leaves open the possibility that there might exist a (possibly far larger) counterexample for powers of 2, that has not yet been detected.

This discussion shows that Conjecture 1.3, if true, implies when for powers of 2 a kind of very strong arithmetic independence of powers of 2 and powers of 3. This would represent an important feature of the prime 2 in integer complexity. Conjecture 1.3 has implications about the number of nonzero digits in the expansion of 2^n in base 3 as a function of n; namely, if there existed a large power of 2 with a huge number of zero digits in its base 3 expansion, then this would give a (counter)-example achieving $\|2^k\| < 2k$. Problems similar to this very special subproblem already appear difficult (see Lagarias [8]). A result of C. L. Stewart[12] yields a lower bound on the number of nonzero digits appearing in the base 3 expansion a power of 2, but it is tiny.

The truth of $\|2^n\| = 2n$ would also immediately imply the lower bound

$$\limsup_{n\to\infty} \frac{\|n\|}{\log n} \geq \frac{2}{\log 2}.$$

Computer experiments seem to agree with this prediction and even allow the possibility of equality, see Iraids et al [7].

There remain many interesting open questions concerning the classification of integers given by the defect. The first concerns the distribution of stable and unstable integers. How many are there of each kind? A second question concerns the function $M(n)$ that counts the number of distinct minimal decompositions into 1's that given integer n has. How does this function behave?

Finally we remark that the set $\mathscr{D} := \{\delta(n) : n \geq 1\}$ of all defect values turns out to be a highly structured set. In a sequel [2], we shall show that it is a well-ordered set, of order type ω^ω, a fact related to some earlier conjectures of Juan Arias de Reyna [4].

2. Properties of the defect

The defect is the fundamental tool in this paper; let us begin by noting some of its basic properties.

Proposition 2.1. *(1) For all integers $a \geq 1$,*

$$\delta(a) \geq 0.$$

Here equality holds for $a = 3^k$, $k \geq 1$.

(2) One has

$$\delta(ab) \leq \delta(a) + \delta(b),$$

and equality holds if and only if $\|ab\| = \|a\| + \|b\|$.

(3) For $k \geq 1$,

$$\delta(3^k \cdot n) \leq \delta(n)$$

and equality holds if and only if $\|3^k \cdot n\| = 3k + \|n\|$.

Proof. (1) This follows from the result of Selfridge. Since for $k \geq 1$, $\|3^k\| = 3k$, we have $\delta(3^k) = 0$ for $k \geq 1$, while $\delta(1) = 1$.

(2) This is a direct consequence of the definition.

(3) This follows from (2), from noting that $\delta(3^k) = 0$ for $k \geq 1$. ∎

Because $\|3^k\| = 3k$ for $k \geq 1$, one might hope that in general, $\|3n\| = 3 + \|n\|$ for $n > 1$. However, this is not so; for instance, $\|107\| = 16$, but $\|321\| = 18$.

The defect measures how far a given integer is from the upper bound $E(||n||)$, given in terms of the ratio $E(\|n\|)/n$:

Proposition 2.2. *We have $\delta(1) = 1$ and*

$$\delta(n) = \begin{cases} 3\log_3 \frac{E(\|n\|)}{n} & \text{if} \quad \|n\| \equiv 0 \pmod 3, \\ 3\log_3 \frac{E(\|n\|)}{n} + 2\,\delta(2) & \text{if} \quad \|n\| \equiv 1 \pmod 3, \text{ with } n > 1, \\ 3\log_3 \frac{E(\|n\|)}{n} + \delta(2) & \text{if} \quad \|n\| \equiv 2 \pmod 3. \end{cases}$$

In particular $E(\|n\|)/n \geq 1$ for any $n \geq 1$.

Proof. The proof is a straightforward computation using Selfridge's formulas for $E(k)$, for $k = 3j + i$, $i = 0, 1, 2$. ∎

2.1. **Stable Integers.** This example above motivates the following definition.

Definition 2.3. A number m is called *stable* if $\|3^k \cdot m\| = 3k + \|m\|$ holds for every $k \geq 1$. Otherwise it is called *unstable*.

We have the following criterion for stability.

Proposition 2.4. *The number m is stable if and only if $\delta(3^k \cdot m) = \delta(m)$ for all $k \geq 0$.*

Proof. This is immediate from Proposition 2.1(3). ∎

These results already suffice to prove the following result, conjectured by Juan Arias de Reyna [4].

Theorem 2.5. *(1) For any $m \geq 1$, there exists a finite $K \geq 0$ such that $3^K m$ is stable.*

(2) If the defect $\delta(m)$ satisfies $0 \leq \delta(m) < 1$, then m itself is stable.

Proof of Theorem 2.5. (1) From Proposition 2.1, we have that for any n, $\delta(3n) \leq \delta(n)$, with equality if and only if $\|3n\| = \|n\| + 3$. More generally, $\delta(3n) = \delta(n) - (\|n\| + 3 - \|3n\|)$, and so the difference $\delta(n) - \delta(3n)$ is always an integer. This means that the sequence $\delta(m), \delta(3m), \delta(9m), \ldots$ is non-increasing, nonnegative, and can only decrease in integral amounts; hence it must eventually stabilize. Applying Proposition 2.4 proves the theorem.

(2) If $\delta(m) < 1$, since all $\delta(n) \geq 0$ there is no room to remove any integral amount, so m must be stable. ∎

Note that while this proof shows that for any n there exists K such that 3^K is stable, it yields no upper bound on such a K. We will give a more constructive proof and obtain an upper bound in a sequel paper [3].

The value of the defect separates the integers into small classes, whose members differ only by powers of 3.

Proposition 2.6. *Suppose that m and n are two positive integers, with $m > n$.*

(1) If $q := \delta(n) - \delta(m)$ is rational, then it is necessarily a nonnegative integer, and furthermore $m = n \cdot 3^k$ for some $k \geq 1$.

(2) If $\delta(n) = \delta(m)$ then $m = n \cdot 3^k$ for some $k \geq 1$ and furthermore

$$||n \cdot 3^j|| = 3j + ||n|| : 0 \leq j \leq k.$$

In particular $\delta(n) = \delta(m)$ implies $\|n\| \equiv \|m\| \pmod 3$.

Proof. (1) If $q = \delta(n) - \delta(m)$ is rational, then $k = \log_3(m/n) \in \mathbb{Q}$ is rational; since m/n is rational, the only way this can occur is if $\log_3(m/n)$ is an integer k, in which case, since $m > n$, $m = n \cdot 3^k$ with $k \geq 1$. It then follows from the definition of defect that $q = \|n\| + 3k - \|m\|$.

(2) By (1) we know that $m = n \cdot 3^k$ for some $k \geq 1$. By Proposition 2.1 (3) we have $\delta(n \cdot 3^j) \leq \delta(n)$, for $j \geq 0$ and it also gives $\delta(m) = \delta(n \cdot 3^k) \leq \delta(n \cdot 3^j)$, for $0 \leq j \leq k$. Since $\delta(m) = \delta(n)$ by hypothesis, this gives $\delta(n \cdot 3^j) = \delta(n)$, so that $||n \cdot 3^j|| = 3j + ||n|| : 0 \leq j \leq k$. $\square$

The results so far suffice to prove Theorem 1.5.

Proof of Theorem 1.5. (1) This follows from Proposition 2.6(2).

(2) The non-increasing assertion follows from Proposition 2.1(3). The finiteness of the number of leaders in a sequence $3^k \cdot n$ follows from Theorem 2.5 (1). $\square$

2.2. **Leaders.** Again because $\|3n\|$ is not always equal to $3 + \|n\|$, it makes sense to introduce the following definition:

Definition 2.7. We call a natural number n a *leader* if it cannot be written most-efficiently as $3m$ for some m; i.e., if either $3 \nmid n$, or, if $3 \mid n$, then $\|n\| < 3 + \|n/3\|$.

For example, 107 is a leader since $3 \nmid 107$, and 321 is also a leader since $\|321\| = 18 < 3 + 16 = 3 + \|107\|$. However, 963 is not a leader, as $\|963\| = 21 = 3 + \|321\|$. Leaders can be stable or unstable. In this example 107 is unstable, but by Theorem 2.5 some multiple $3^K \cdot 107$ will be stable, and the smallest such multiple will be a stable leader.

We have the following alternate characterization of leaders:

Proposition 2.8. *(1) A number n is a leader if and only if it is the smallest number having its given defect value.*

(2) For any natural number m, there is a unique leader $n \leq m$ such that $\delta(n) = \delta(m)$. For it $m = n \cdot 3^k$ for some $k \geq 0$.

Proof. (1) If this were false, there would a leader n with some $n' < n$ with $\delta(n') = \delta(n)$. By Proposition 2.6 (2) $n = 3^k \cdot n'$ with $k \geq 1$ and $||n' \cdot 3^j|| = 3j + ||n'||$ for $0 \leq j \leq k$. But then $n/3 = n' \cdot 3^{k-1}$ is an integer and $||n/3|| = ||n'|| + 3k = ||n|| - 3$, which contradicts n being a leader.

Conversely, if n is the first number of its defect and is divisible by 3, then we cannot have $\|n\| = \|n/3\| + 3$, or else by Proposition 2.1 we would obtain $\delta(n) = \delta(n/3)$, contradicting minimality.

(2) Pick n to be the smallest number such that $\delta(n) = \delta(m)$; this is the unique leader satisfying $\delta(n) = \delta(m)$. Then $m = 3^k n$ for some $k \geq 0$ by Proposition 2.6. $\square$

To summarize, if δ occurs as a defect, then the set of integers

$$N(\delta) := \{m : \delta(m) = \delta\},$$

having a given defect value δ has a smallest element that is a leader. If this leader n is unstable, then this set of integers $N(\delta) = \{3^j \cdot n : 0 \leq j \leq j(\delta)\}$. If this leader is stable, then $N(\delta) = \{3^j \cdot n : j \geq 0\}$ is an infinite set. Furthermore if $3 \nmid n$ then n is a leader, and there is a unique $K = K(n) \geq 0$ such that $n' = 3^K n$ is a stable leader.

3. Good factorizations and solid numbers

Given a natural number $n > 1$, by the dynamic programming definition of complexity there are either two numbers u and v, both smaller than n, such that $n = u \cdot v$ and $\|n\| = \|u\| + \|v\|$, or such that $n = u + v$ and $\|n\| = \|u\| + \|v\|$. In the case u and v such that $n = u + v$, and $\|n\| = \|u\| + \|v\|$ we say n is *additively reducible*. In the case $n = u \cdot v$ and $\|n\| = \|u\| + \|v\|$ we say n is *multiplicatively reducible*. Some numbers n are reducible in both senses. For instance, $10 = 9 + 1$ with $\|10\| = \|9\| + \|1\|$, and $10 = 2 \cdot 5$ with $10 = \|2\| + \|5\|$.

We call any such (additive or multiplicative) decomposition preserving equality of complexities a *most-efficient decomposition*. In the additive case this is a partitioning of n in two parts while in the multiplicative case it is a factorization of n in two factors.

3.1. **Additive Irreducibility and Solid Numbers.**

We introduce terminology for numbers not being additively reducible.

Definition 3.1. We will say that a natural number n is *additively irreducible* if it cannot be written most efficiently as a sum, i.e., for all u and v such that $n = u+v$, we have $\|n\| < \|u\| + \|v\|$. We call such values of n *solid numbers*.

The first few solid numbers are

$$\{1, 6, 8, 9, 12, 14, 15, 16, 18, 20, 21, 24, 26, 27, \ldots...\}$$

Experimental evidence suggests that a positive fraction of integers below x are solid numbers, as $x \to \infty$.

3.2. **Multiplicative Irreducibility and Good Factorizations.**

We introduce further terminology for multiplicative factorizations that respect the complexity.

Definition 3.2. A factorization $n = u_1 \cdot u_2 \cdots u_k$ is a *good factorization* of n if the following equality holds:

$$\|n\| = \|u_1\| + \|u_2\| + \ldots + \|u_k\|.$$

The factorization containing only one factor is automatically good; this will be called a *trivial good factorization*.

Proposition 3.3. *If $n = n_1 \cdot n_2 \cdot \ldots \cdot n_k$ is a good factorization then for any nonempty subset $I \subset \{1, 2, \ldots, k\}$ the product $m = \prod_{j \in I} n_j$ is a good factorization of m.*

Proof. If the factorization of m were not good, then we would have

$$\|m\| < \sum_{j \in I} \|n_j\|$$

But then

$$\|n\| = \left\| m \prod_{j \notin I} n_j \right\| < \sum_{j \in I} \|n_j\| + \sum_{j \notin I} \|n_j\| = \sum_{j=1}^{k} \|n_j\|$$

and the given factorization of n would not be a good factorization. $\square$

Proposition 3.4. *(1) If $n = n_1 \cdot n_2 \cdot ... \cdot n_k$ is a good factorization, and each $n_i = n_{i,1} \cdot \ldots \cdot n_{i,l_i}$ is a good factorizations, then so is $n = \prod_{i=1}^{k} \prod_{j=1}^{l_i} n_{i,j}$.*

(2) If $n = n_1 \cdot n_2 \cdot \ldots \cdot n_k$ is a good factorization, and $I_1, I_2, \ldots, I_l$ is a partition of $\{1, \ldots, k\}$, then letting $m_i = \prod_{j \in I_i} n_j$, we have that $n = \prod_{i=1}^{l} m_i$ is a good factorization.

Proof. (1)We have that $\|n_i\| = \sum_{j=1}^{l_i} \|n_{i,j}\|$ and $\|n\| = \sum_{i=1}^{k} \|n_i\|$, so

$$\|n\| = \sum_{i=1}^{k} \sum_{j=1}^{l_i} \|n_{i,j}\|$$

and we are done.

(2) This follows from Proposition 3.3 together with (1). □

Definition 3.5. We will say that a natural number n is *multiplicatively irreducible* (abbreviated *m-irreducible*) if n has no nontrivial good factorizations.

Proposition 3.4(2) shows n is m-irreducible if and only if all nontrivial factorizations $n = uv$ have $\|n\| < \|u\| + \|v\|$. Thus a prime number p is automatically m-irreducible since the only factorization is $p = p \cdot 1$ and obviously we have $\|p\| < \|p\| + 1 = \|p\| + \|1\|$. However, the converse does not hold. For instance, 46 is a composite number which is m-irreducible.

Proposition 3.6. *Any natural number has a good factorization into m-irreducibles.*

Proof. We may apply induction and assume that any $m < n$ has a factorization in m-irreducibles. If n is m-irreducible, we are done. Otherwise, n has a good factorization $n = uv$. Observe that $n = n \cdot 1$ is never a good factorization, since $\|1\| = 1$; hence, u, $v < n$. Then the induction hypothesis implies that there are good factorizations for u and v. Multiplying them together and applying Proposition 3.4, we obtain a good factorization of n. □

Good factorizations into m-irreducibles need not be unique. For $4838 = 2 \cdot 41 \cdot 59$, we find that $2 \cdot (41 \cdot 59)$, $(2 \cdot 59) \cdot 41$ and $(2 \cdot 41) \cdot 59$ are all good factorizations, but the full factorization $2 \cdot 41 \cdot 59$ is not a good factorization. (Thanks to Juan Arias de Reyna for this example.) This is deducible from the following data:

$$\|2 \cdot 41 \cdot 59\| = 27,$$
$$\|2\| = 2, \quad \|41\| = 12, \quad \|59\| = 14.$$
$$\|2 \cdot 41\| = 13, \quad \|2 \cdot 59\| = 15, \quad \|41 \cdot 59\| = 25,$$

3.3. **Good factorizations and leaders.** The next two propositions show how the notion of good factorization interacts with leaders and stability.

Proposition 3.7. *Let $n = n_1 \cdot n_2 \cdots n_r$ be a good factorization. If n is a leader then each of the factors n_j is a leader.*

Proof. Suppose otherwise; without loss of generality, we may assume that n_1 is not a leader, so $3 \mid n_1$ and $\|n_1\| = 3 + \|n_1/3\|$. So $3 \mid n$ and

$$\|n/3\| = \|(n_1/3) \cdot n_2 \cdot \ldots \cdot n_r\| \leq \|n_1/3\| + \sum_{j=2}^{r} \|n_j\| =$$

$$= \|n_1\| - 3 + \sum_{j=2}^{r} \|n_j\| = \|n\| - 3.$$

Since $\|n\| \leq 3 + \|n/3\|$, we have $\|n\| = 3 + \|n/3\|$, thus n is not a leader. $\square$

Proposition 3.8. *Let $n = n_1 \cdot n_2 \cdots n_r$ be a good factorization. If n is stable, then each of its factors n_j is stable.*

Proof. Suppose otherwise. Without loss of generality, we may assume that n_1 is unstable; say $\|3^k n_1\| < \|n_1\| + 3k$. So

$$\|3^k n\| = \|(3^k n_1) \cdot n_2 \cdot \ldots \cdot n_r\| \leq \|3^k n_1\| + \sum_{j=2}^{r} \|n_j\| =$$

$$< \|n_1\| + 3k + \sum_{j=2}^{r} \|n_j\| = \|n\| + 3k.$$

and thus n is not stable. $\square$

Assembling all these results we deduce that being a leader and being stable are both inherited properties for subfactorizations of good factorizations.

Proposition 3.9. *Let $n = n_1 \cdot n_2 \cdots n_r$ be a good factorization, and I be a nonempty subset of $\{1, \ldots, r\}$; let $m = \prod_{i \in I} n_i$. If n is a leader, then so is m. If n is stable, then so is m.*

Proof. Immediate for Proposition 3.7, Proposition 3.8, and Proposition 3.4.(2). $\square$

4. The Classification Method

Here, we state and prove a result (Theorem 4.4) that will be our primary tool for the rest of the paper. By applying it repeatedly, for any $r > 0$, we can put restrictions on what integers n can satisfy $\delta(n) < r$.

Definition 4.1. (1) For any real $r \geq 0$, define A_r to be $\{n \in \mathbb{N} : \delta(n) < r\}$.
(2) Define B_r to be the set consisting of those elements of A_r that are leaders.

While A_r is our main object of interest, it turns out to be easier and more natural to deal with B_r. Note that knowing B_r is enough to determine A_r, as expressed in the following proposition:

Proposition 4.2.

$$A_r = \{3^k n : n \in B_r, k \geq 0\}$$

Proof. If $n \in B_r$, then $\delta(3^k n) \leq \delta(n) < r$, so $3^k n \in A_r$. Conversely, if $m \in A_r$, by Proposition 2.8(2) we can take $n \geq 1$ and $k \geq 0$ such that n is a leader, $m = 3^k n$, and $\delta(m) = \delta(n)$; then $n \in B_r$ and we are done. $\square$

We now let $\alpha > 0$ be a real parameter, specifiable in advance. The main result puts constraints on the allowable forms of the dynamic programming recursion (most efficient representations) to compute integers in $B_{(k+1)\alpha}$ in terms of integers in $B_{j\alpha}$ for $1 \le j \le k$. However there are some exceptional cases where the theorem does not work very well; fortunately, for any $\alpha < 1$, there are only finitely many. We will collect these into a set we call T_α.

Definition 4.3. Define T_α to consist of 1 together with those m-irreducible numbers n which satisfy

$$\frac{1}{n-1} > 3^{\frac{1-\alpha}{3}} - 1$$

and do not satisfy $\|n\| = \|n-b\| + \|b\|$ for any solid numbers b with $1 < b \le n/2$.

Observe that for $0 < \alpha < 1$, the above inequality is equivalent to

$$n < (3^{\frac{1-\alpha}{3}} - 1)^{-1} + 1$$

and hence T_α is a finite set. For $\alpha \ge 1$, the inequality is trivially satisfied and so $T_\alpha = T_1$. We do not know whether T_1 is a finite or an infinite set. However in our computations we will always choose values $0 < \alpha < 1$.

We can now state the main classification result, which puts strong constraints on the form of most efficient decompositions on numbers in sets $B_{(k+1)\alpha}$.

Theorem 4.4. *Suppose $0 < \alpha < 1$ and that $k \ge 1$. Then any $n \in B_{(k+1)\alpha}$ can be most-efficiently represented in (at least) one of the following forms:*

(1) *For $k = 1$, there is either a good factorization $n = u \cdot v$ where $u, v \in B_\alpha$, or a good factorization $n = u \cdot v \cdot w$ with $u, v, w \in B_\alpha$;*
For $k \ge 2$, there is a good factorization $n = u \cdot v$ where $u \in B_{i\alpha}$, $v \in B_{j\alpha}$ with $i + j = k + 2$ and $2 \le i, j \le k$.
(2) *$n = a + b$ with $\|n\| = \|a\| + \|b\|$, $a \in A_{k\alpha}$, $b \le a$ a solid number and*
$$\delta(a) + \|b\| < (k+1)\alpha + 3\log_3 2.$$
(3) *There is a good factorization $n = (a+b)v$ with $v \in B_\alpha$ and a and b satisfying the conditions in the case (2) above.*
(4) *$n \in T_\alpha$ (and thus in particular either $n = 1$ or $\|n\| = \|n-1\| + 1$.)*
(5) *There is a good factorization $n = u \cdot v$ with $u \in T_\alpha$ and $v \in B_\alpha$.*

We will prove Theorem 4.4 in Section 4.2, after establishing a preliminary combinatorial lemma in Section 4.1.

To apply Theorem 4.4, one recursively constructs from given sets $B^*_{j\alpha}$, $A^*_{j\alpha}$ for $1 \le j \le k-1$ which contain $B_{j\alpha}, A_{j\alpha}$, respectively, the set of all n satisfying the relaxed conditions (1)-(5) obtained replacing $B_{j\alpha}$ by $B^*_{j\alpha}$ and $A_{j\alpha}$ by $A^*_{j\alpha}$. This new set $B^{**}_{(k+1)\alpha}$ contains the set $B_{(k+1)\alpha}$ we want. Sometimes we can, by other methods, prune some elements from $B^{**}_{(k+1)\alpha}$ that do not belong to $B_{(k+1)\alpha}$, to obtain a new approximation B^*_{k+1}. This then determines $A^*_{(k+1)\alpha} := \{3^k n :\ k \ge 0, n \in B^*_{(k+1)\alpha}\}$, permitting continuation to the next level $k+2$. We will present two applications of this construction:

(1) To get an upper bound on the cardinality of $B_{(k+1)\alpha}$ of numbers below a given bound x.
(2) To get a lower bound for the complexity $\|n\|$ of a number n by showing it does not belong to a given set $B^*_{k\alpha}$; this excludes it from $B_{k\alpha}$, whence $\|n\| \ge 3\log_3 n + k\alpha$.

In some circumstances we can obtain the exact sets $B_{k\alpha}$ and $A_{k\alpha}$ for $1 \leq k \leq k_0$, i.e. we recursively construct $B^*_{k\alpha}$ so that $B^*_{k\alpha} = B_{k\alpha}$. This requires a perfect pruning operation at each step. Here a good choice of the parameter α is helpful.

In applications we will typically not use the full strength of Theorem 4.4. Though the representations it yields are most efficient, the proofs will typically not use this fact. Also, in the addition case (2), the requirement that $\delta(a) + \|b\| < (k+1)\alpha + 3\log_3 2$ implies the weaker requirement that just $\|b\| < (k+1)\alpha + 3\log_3 2$. The latter relaxed condition is easier to check, but it does enlarge the initial set $B^{**}_{(k+1)\delta}$ to be pruned.

4.1. **A Combinatorial Lemma.** We establish a combinatorial lemma regarding decomposing a sum of real numbers into blocks.

Lemma 4.5. *Let $x_1, x_2, \ldots, x_r > 0$ be real numbers such that $\sum_{i=1}^r x_i < k+1$, where $k \geq 1$ is a natural number.*

(1) If $k \geq 2$ then either there is some i with $x_i \geq k$, or else we may find a partition $A \cup B$ of the set $\{1, 2, \ldots, r\}$ such that

$$\sum_{i\in A} x_i < k, \qquad \sum_{i \in B} x_i < k.$$

(2) If $k = 1$ then either there is some i with $x_i \geq 1$, or else we may find a partition $A \cup B \cup C$ of the set $\{1, 2, \ldots, r\}$ such that

$$\sum_{i\in A} x_i < 1, \qquad \sum_{i\in B} x_i < 1, \qquad \sum_{i\in C} x_i < 1.$$

Proof. (1) Suppose $k \geq 2$. Let us abbreviate $\sum_{i\in S} x_i$ by $\sum S$. Among all partitions $A \cup B$ of $\{1, \ldots, r\}$, take one that minimizes $|\sum A - \sum B|$, with $\sum A \geq \sum B$. Suppose that $\sum A \geq k$; then since $\sum A + \sum B < k+1$, we have $\sum B < 1$, and so $\sum A - \sum B > k-1$. So pick $x_i \in A$ and let $A' = A \setminus \{i\}$, $B' = B \cup \{i\}$. If $\sum A' > \sum B'$, then $|\sum A' - \sum B'| = \sum A - \sum B - 2x_i < \sum A - \sum B$, contradicting minimality, so $\sum A' \leq \sum B'$. So $\sum B' - \sum A' \geq \sum A - \sum B$, i.e.,

$$x_i \geq \sum A - \sum B > k-1.$$

Now i was an arbitrary element of A; this means that A can have at most one element, since otherwise, if $j \neq i \in A$, we would have $\sum A \geq x_i + x_j$ and hence $x_j \leq \sum B < 1$, but also $x_j > k-1$, contradicting $k \geq 2$. Thus $A = \{i\}$ and so $x_i \geq k$.

(2) Here $k = 1$. Assume that $x_1 \geq x_2 \geq \cdots \geq x_r$. If $x_1 \geq 1$ we are done. Otherwise, if $r \leq 3$, we can partition $\{1, \ldots, r\}$ into singletons.

For $r \geq 4$, assume by induction the lemma is true for all sets of numbers with strictly less than r elements. Let $y = x_{r-1} + x_r$. We must have $y < 1$ because otherwise $x_{r-3} + x_{r-2} \geq x_{r-1} + x_r \geq 1$ and we get $\sum_{i=1}^r x_i \geq 2$ in contradiction to the hypothesis. Hence, if we define $x'_1 = x_1, \ldots, x'_{r-2} = x_{r-2}$, $x'_{r-1} = y$, we have $\sum_{i=1}^{r-1} x'_i = \sum_{i=1}^r x_i < 2$, and $x'_i < 1$ for all i. By the inductive hypothesis, then, there exists a paritition $A' \cup B' \cup C' = \{1, \ldots, r-1\}$ with

$$\sum_{i\in A'} x'_i < 1, \quad \sum_{i\in B'} x'_i < 1, \quad \sum_{i\in C'} x'_i < 1.$$

Replacing x'_{r-1} with x_{r-1} and x_r, we get the required partition of $\{1, \ldots, r\}$. $\square$

For $k=1$ the example taking $\{x_1, x_2, x_3\} = \{3/5, 3/5, 3/5\}$ shows that a partition into three sets is sometimes necessary.

4.2. **Proof of the Classification Method.**

Proof of Theorem 4.4. Suppose $n \in B_{(k+1)\alpha}$; take a most-efficient representation of n, which is either ab, $a+b$, or 1. If $n=1$, then $n \in T_\alpha$ and we are in case (4). So suppose $n>1$.

If n is m-irreducible, we will pick a way of writing $n=a+b$ with $\|n\| = \|a\|+\|b\|$, $a \geq b$, and b is solid. There is necessarily a way to do this, since ne way to do so is to write $n=a+b$ with $\|n\| = \|a\|+\|b\|$ and b minimal. Since this is possible, then, if there is a way to choose a and b to have $b>1$, do so; otherwise, we must pick $b=1$. In either case,

$$\|a\|+\|b\| = \|n\| < 3\log_3(a+b) + (k+1)\alpha \leq 3\log_3(2a) + (k+1)\alpha,$$

so $\delta(a) + \|b\| < (k+1)\alpha + 3\log_3 2$.

If $a \in A_{k\alpha}$, we are in case (2). Otherwise, we have

$$3\log_3 a + k\alpha + \|b\| \leq \|a\| + \|b\| = \|n\| <$$
$$3\log_3(a+b) + (k+1)\alpha \leq 3\log_3(2a) + (k+1)\alpha,$$

so $\|b\| < 3\log_3 2 + \alpha$; since $\alpha < 1$, we have $\|b\| \leq 2$ and thus $b \leq 2$; because b is solid, we have $b=1$. By assumption, we only picked $b=1$ if this choice was forced upon us, so in this case, we must have that n does not satisfy $\|n\| = \|n-b\| + \|b\|$ for any solid b with $1 < b \leq b/2$.

Since $\alpha < 1$ we have $3\log_3 a + k\alpha + 1 < 3\log_3(a+1) + (k+1)\alpha$; if we solve for a, we find that

$$\frac{1}{n-1} = \frac{1}{a} > 3^{\frac{1-\alpha}{3}} - 1.$$

Thus, $n \in T_\alpha$ and we are in case (4).

Now we consider the case when n is not m-irreducible. Choose a good factorization of n into m-irreducible numbers; since n is not m-irreducible, $n = \prod_{i=1}^r m_i$, with $r \geq 2$. Then we have $\sum_{i=1}^r \delta(m_i) = \delta(n) < (k+1)\alpha$. Note that since we assumed n is a leader, every product of a proper subset of the m_i is also a leader by Proposition 3.9. We now have two cases.

Case 1. $k \geq 2$.

Now by Lemma 4.5(1), either there exists an i with $\delta(m_i) \geq k\alpha$, or else we can partition the $\delta(m_i)$ into two sets each with sum less than $k\alpha$.

In the latter case, we may also assume these sets are nonempty, as if one is empty, this implies that $\delta(n) < k\alpha$, and hence any partition of the $\delta(m_i)$ will work; since $r \geq 2$, we can take both these sets to be nonempty. In this case, call the products of these two sets u and v, so that $n = uv$ is a good factorization of n. Then $\delta(u) + \delta(v) < (k+1)\alpha$, so if we let $(i-1)\alpha$ be the largest integral multiple of α which is at most $\delta(u)$, then letting $j = k+2-i$, we have $\delta(v) < j\alpha$. So $i+j = k+2$; furthermore, since $i\alpha$ is the smallest integral multiple of α which is greater than $\delta(u)$, and $\delta(u) \leq \delta(n) < k\alpha$, we have $i \leq k$, so $j \geq 2$. If also $i \geq 2$ then $j \leq k$, and so we are in case (1). If instead $i=1$, then we have $u \in B_\alpha \subseteq B_{2\alpha}$, and $v \in B_{k\alpha}$ (since $\delta(v) \leq \delta(n) < k\alpha$), so we are again in case (1) if we take $i=2$ and $j=k$.

If such a partition is not possible, then let u be an m_i with $\delta(m_i) \geq k\alpha$, and let v be the product of the other m_i, so that once again $n = uv$ is a good factorization of n. Since $\delta(u)+\delta(v)=\delta(n)$, we have $\delta(v)<\alpha$, and so $v\in B_\alpha$. Finally, since u is m-irreducible and an element of $B_{(k+1)\alpha}$, it satisfies the conditions of either case (2) or case (4), and so n satisfies the conditions of either case (3) or case (5).

Case 2. $k=1$.

Now by Lemma 4.5(2), either there exists an i with $\delta(m_i)\ge\alpha$, or else we can partition the $\delta(m_i)$ into three sets each with sum less than α.

In the latter case, we may also assume at least two of these sets are nonempty, as otherwise $\delta(n)<\alpha$, and hence any partition of the $\delta(m_i)$ will work. If there are two nonempty sets, call the products of these two sets u and v, so that $n=uv$ is a good factorization of n. If there are three nonempty sets, call them u,v,w, so that $n=uvw$ is a good factorization of n. Thus we are in case (1) for $k=1$.

If such a partition is not possible, then we repeat the argument in Case 1 above, determining that n satisfies one of the conditions of cases (3) or (5). □

5. Determination of all elements of defect below a given bound r

In this section we determine all elements of A_r for certain small r, using Theorem 4.4 together with a pruning operation.

5.1. Classification of numbers of small defect.

We will now choose as our parameter

$$\alpha := \delta(2) = 2 - 3\log_3 2 \approx 0.107.$$

The choice of this parameter is motivated by Theorem 5.2 below. We use above method to inductively compute $A_{k\delta(2)}$ and $B_{k\delta(2)}$ for $0\le k\le 12$. Numerically, $1.286<12\delta(2)<1.287$. The following result classifies all integers in $A_{12\delta(2)}$.

Theorem 5.1. (Classification Theorem) *The numbers n satisfying $\delta(n)<12\delta(2)$ are precisely those that can be written in at least one of the following forms, which have the indicated complexities:*

1. 3^k *of complexity* $3k$ *(for $k\ge 1$)*
2. 2^a3^k *for $a\le 11$, of complexity* $2a+3k$ *(for a, k not both zero)*
3. $5\cdot 2^a3^k$ *for $a\le 6$, of complexity* $5+2a+3k$
4. $7\cdot 2^a3^k$ *for $a\le 5$, of complexity* $6+2a+3k$
5. $19\cdot 2^a3^k$ *for $a\le 3$, of complexity* $9+2a+3k$
6. $13\cdot 2^a3^k$ *for $a\le 2$, of complexity* $8+2a+3k$
7. $2^a(2^b3^l+1)3^k$ *for $a+b\le 2$, of complexity* $2(a+b)+3(l+k)+1$ *(for b, l not both zero).*
8. 1, *of complexity* 1
9. $55\cdot 2^a3^k$ *for $a\le 2$, of complexity* $12+2a+3k$
10. $37\cdot 2^a3^k$ *for $a\le 1$, of complexity* $11+2a+3k$
11. $25\cdot 3^k$ *of complexity* $10+3k$
12. $17\cdot 3^k$ *of complexity* $9+3k$
13. $73\cdot 3^k$ *of complexity* $13+3k$

This list is redundant; for example list (7) with $a=0,b=1,l=1$ gives $7\cdot 3^k$, which overlaps list (4) with $a=0$. But the given form is convenient for later purposes. In the next section we will give several applications of this result. They can be derived knowing only the statement of this theorem, without its proof, though one will also require Theorem 4.4.

The detailed proof of this theorem is given in the rest of this section. The proof recursively determines all the sets $A_{k\delta(2)}$ and $B_{k\delta(2)}$ for $1\le k\le 12$. It is possible to extend this method to values $n\delta(2)$ with $n > 12$ but it is tedious. In a sequel paper [3], we will present a method for automating these computations.

5.2. **Base case.** The use of $\delta(2)$ may initially seem like an odd choice of step size. Its significance is shown by the following base case, which is proved using Rawsthorne's result that $E_1(n) \leq (8/9)E(n)$ (with equality for $n \geq 8$).

Theorem 5.2. *If $\delta(n) \neq 0$, then $\delta(n) \geq \delta(2)$. Equivalently, if n is not a power of 3, then $\delta(n) \geq \delta(2)$.*

Proof. We apply Proposition 2.2. There are four cases

Case 1. If $n = 1$, then $\delta(n) = 1 \geq \delta(2)$.

Case 2. If $\|n\| \equiv 2 \pmod 3$, then

$$\delta(n) = \delta(2) + 3\log_3 \frac{E(\|n\|)}{n} \geq \delta(2).$$

Case 3. If $\|n\| \equiv 1 \pmod 3$ and $n > 1$, then

$$\delta(n) = 2\delta(2) + 3\log_3 \frac{E(\|n\|)}{n} \geq 2\delta(2) \geq \delta(2).$$

Case 4. If $\|n\| \equiv 0 \pmod 3$, then $\delta(n) = 3\log_3(E(\|n\|)/n)$. We know that in this case $n = E(\|n\|)$ if and only if n is a power of 3 if and only if $\delta(n) = 0$. So if $\delta(n) \neq 0$, then $n \leq E_1(\|n\|)$. But $E_1(\|n\|) \leq (8/9)E(\|n\|)$, so $E(\|n\|)/n \geq 9/8$, so $\delta(n) \geq 3\log_3 \frac{9}{8} = 3\delta(2) \geq \delta(2)$. ☐

The proof above also established:

Proposition 5.3. *$B_0 = \emptyset$, and $B_{\delta(2)} = \{3\}$.*

To prove Theorem 5.1 we will use Theorem 4.4 for the "inductive step". However, while Theorem 4.4 allows us to place restrictions on what A_r can contain, if we want to determine A_r itself, we need a way to certify membership in it. To certify inclusion in A_r we need an upper bound on the defect, which translates to an upper bound on complexity, which is relatively easy to do. However we also need to discard n that do not belong to A_r, i.e. pruning the set we are starting with. This requires establishing lower bounds on their defects, certifying they are r or larger, and for this we need lower bounds on their complexities.

5.3. **Two pruning lemmas.** To find lower bounds on complexities, we typically use the following technique. Say we want to show that $\|n\| \geq k$ ($k \in \mathbb{N}$); since $\|n\|$ is always an integer, it suffices to show $\|n\| > k - 1$. We do this by using our current knowledge of A_l for various l; by showing that if $\|n\| \leq k - 1$ held, then it would put n in some A_l which we have already determined and know it's not in. The following two lemmas, both cases of this principle, are useful for this purpose.

Lemma 5.4. *Take $\alpha \leq 1/2$. Say $\delta(a) < i\alpha$ and $\delta(b) < j\alpha$, and let $k + 2 = i + j$. Then $\|ab\| = \|a\| + \|b\|$ holds unless $\delta(ab) < k\alpha$.*

Proof. Note

$$\|ab\| \geq 3\log_3(ab) + k\alpha = 3\log_3 a + 3\log_3 b + (i + j - 2)\alpha > \|a\| + \|b\| - 1$$

so $\|ab\| \geq \|a\| + \|b\|$. ☐

Lemma 5.5. *If $\delta(a) < k\alpha$, then $\|3^m(a+1)\| = 3m + \|a\| + 1$ holds unless $\delta(3^m(a+1)) < k\alpha$.*

Proof. Note

$$\|3^m(a+1)\| \geq 3\log_3(a+1) + 3m + k\alpha > \|a\| + 3m$$

so $\|3^m(a+1)\| \geq 3m + \|a\| + 1$. □

5.4. **Proof of Theorem 5.1: Inductive Steps.** We prove Theorem 5.1 by repeatedly applying Theorem 4.4, to go from k to $k+1$ for $0 \leq k \leq 12$. We will use a step size $\alpha = \delta(2)$, so let us first determine $T_{\delta(2)}$. We compute that $3 < (3^{\frac{1-\delta(2)}{3}} - 1)^{-1} + 1 < 4$, and so $T_{\delta(2)} = \{1, 2, 3\}$. We note that in all cases of attempting to determine $B_{(k+1)\alpha}$ we are considering, we will have $(k+1)\alpha \leq 12\delta(2)$, and so if $\|b\| < (k+1)\alpha + 3\log_3 2$, then

$$\|b\| < 12\delta(2) + 3\log_3 2 = 3.179\ldots,$$

so $\|b\| \leq 3$, which for b solid implies $b = 1$.

The base cases $B_0 = \emptyset$ and $B_{\delta(2)} = \{3\}$ were handled in Proposition 5.3. We now treat the $B_{k\delta}$ in increasing order.

Proposition 5.6.

$$B_{2\delta(2)} = B_{\delta(2)} \cup \{2\},$$

and the elements of $A_{2\delta(2)}$ have the complexities listed in Theorem 5.1.

Proof. By the main theorem,

$$\begin{aligned} B_{2\delta(2)} \setminus B_{\delta(2)} \quad \subseteq \quad & \{1, 2, 6, 9, 27\} \cup \\ & \{3 \cdot 3^n + 1 : n \geq 0\} \cup \{3(3 \cdot 3^n + 1) : n \geq 0\}. \end{aligned}$$

We can exclude 1 because $\delta(1) = 1$, and we can exclude 6, 9, and 27 as they are not leaders. For $3^{n+1} + 1$, Lemma 5.5 shows $\|3^{n+1} + 1\| = 3(n+1) + 1$, and thus $\delta(3^{n+1} + 1) = 1 - 3\log_3(1 + 3^{-(n+1)})$, which allows us to check that none of these lie in $A_{2\delta(2)}$. We can exclude $3(3^{n+1} + 1)$ since Lemma 5.5 shows it has the same defect as $3^{n+1} + 1$ (and so therefore also is not a leader). Finally, checking the complexity of $2 \cdot 3^k$ can be done with Lemma 5.4. □

To make later computations easier, let us observe here that $\delta(3^1 + 1) = \delta(4) = 2\delta(2)$; $6\delta(2) < \delta(3^2 + 1) = \delta(10) < 7\delta(2)$; $8\delta(2) < \delta(3^3 + 1) = \delta(28) < 9\delta(2)$; and that for $n \geq 4$, $9\delta(2) < \delta(3^n + 1) < 10\delta(2)$.

In the above, for illustration, we explicitly considered and excluded 3, 6, 9, 27, and $3(3^{n+1} + 1)$, but henceforth we will simply not mention any multiplications by 3. If $n = 3a$ is a good factorization, n cannot be a leader (by definition), but if it is not a good factorization, we can by Theorem 4.4 ignore it.

Proposition 5.7.

$$B_{3\delta(2)} = B_{2\delta(2)} \cup \{4\},$$

and the elements of $A_{3\delta(2)}$ have the complexities listed in Theorem 5.1.

Proof. By the main theorem,

$$\begin{aligned} B_{3\delta(2)} \setminus B_{2\delta(2)} \quad \subseteq \quad & \{1, 4\} \cup \\ & \{3 \cdot 3^n + 1 : n \geq 0\} \cup \{2 \cdot 3^n + 1 : n \geq 0\}. \end{aligned}$$

Again, $\delta(1) = 1$. By the above computation, the only number of the form $3^{n+1} + 1$ occuring in $A_{3\delta(2)}$ is 4. Lemma 5.5 shows that $\|2 \cdot 3^n + 1\| = 3 + 3n$ for $n > 0$, and hence $\delta(2 \cdot 3^n + 1) = 3 - 3\log_3(2 + 3^{-n})$, which allows us to check that none of these lie in $A_{3\delta(2)}$. Finally, checking the complexity of $4 \cdot 3^k$ can be done with Lemma 5.4. □

To make later computations easier, let us observe here that $6\delta(2) < \delta(2\cdot3^1+1) = \delta(7) < 7\delta(2)$; $8\delta(2) < \delta(2\cdot3^2+1) = \delta(19) < 9\delta(2)$; $9\delta(2) < \delta(2\cdot3^3+1) = \delta(55) < 10\delta(2)$; and that for $n \geq 4$, $10\delta(2) < \delta(2\cdot3^n+1) < 11\delta(2)$.

We will henceforth stop explicitly considering and then excluding 1, since we know that $9\delta(2) < \delta(1) = 1 < 10\delta(2)$.

Proposition 5.8.

$$B_{4\delta(2)} = B_{3\delta(2)} \cup \{8\},$$

and the elements of $A_{4\delta(2)}$ have the complexities listed in Theorem 5.1.

Proof. By the main theorem,

$$\begin{aligned} B_{4\delta(2)} \setminus B_{3\delta(2)} \quad \subseteq \quad & \{8\} \cup \{3\cdot3^n+1 : n \geq 0\} \cup \\ & \{2\cdot3^n+1 : n \geq 0\} \cup \{4\cdot3^n+1 : n \geq 0\}. \end{aligned}$$

By the above computation, no numbers of the form $3^{n+1}+1$ or $2\cdot3^n+1$ occur in $A_{4\delta(2)} \setminus A_{3\delta(2)}$. Lemma 5.5 shows $\|4\cdot3^n+1\| = 5+3n$ and hence $\delta(4\cdot3^n+1) = 5-3\log_3(4+3^{-n})$, which allows us to check that none of these lie in $A_{4\delta(2)}$. Finally, checking the complexity of $8\cdot3^k$ can be done with Lemma 5.4. □

To make later computations easier, let us observe here that $5\delta(2) < \delta(4\cdot3^0+1) = \delta(5) < 6\delta(2)$; $9\delta(2) < \delta(4\cdot3^1+1) = \delta(13) < 10\delta(2)$; $10\delta(2) < \delta(4\cdot3^2+1) = \delta(37) < 11\delta(2)$; and that for $n \geq 3$, $11\delta(2) < \delta(4\cdot3^n+1) < 12\delta(2)$.

Proposition 5.9.

$$B_{5\delta(2)} = B_{4\delta(2)} \cup \{16\},$$

and the elements of $A_{5\delta(2)}$ have the complexities listed in Theorem 5.1.

Proof. By the main theorem,

$$\begin{aligned} B_{5\delta(2)} \setminus B_{4\delta(2)} \quad \subseteq \quad & \{16\} \cup \{3\cdot3^n+1 : n \geq 0\} \cup \{2\cdot3^n+1 : n \geq 0\} \cup \\ & \{4\cdot3^n+1 : n \geq 0\} \cup \{8\cdot3^n+1 : n \geq 0\}. \end{aligned}$$

By the above computation, no numbers of the form $3^{n+1}+1$, $2\cdot3^n+1$, or $4\cdot3^n+1$ occur in $A_{5\delta(2)} \setminus A_{4\delta(2)}$. Lemma 5.5 shows that $\|8\cdot3^n+1\| = 7+3n$ for $n > 0$, and hence $\delta(8\cdot3^n+1) = 7-3\log_3(8+3^{-n})$, which allows us to check that none of these lie in $A_{5\delta(2)}$. Finally, checking the complexity of $16\cdot3^k$ can be done with Lemma 5.4. □

To make later computations easier, let us observe here that $11\delta(2) < \delta(8\cdot3^1+1) = \delta(25) < \delta(8\cdot3^2+1) = \delta(73) < 12\delta(2)$, and that for $n \geq 3$, $\delta(8\cdot3^n+1) > 12\delta(2)$.

Proposition 5.10.

$$B_{6\delta(2)} = B_{5\delta(2)} \cup \{32, 5\},$$

and the elements of $A_{6\delta(2)}$ have the complexities listed in Theorem 5.1.

Proof. By the main theorem,

$$\begin{aligned} B_{6\delta(2)} \setminus B_{5\delta(2)} \quad \subseteq \quad & \{32\} \cup \{3\cdot3^n+1 : n \geq 0\} \cup \{2\cdot3^n+1 : n \geq 0\} \cup \\ & \{4\cdot3^n+1 : n \geq 0\} \cup \{8\cdot3^n+1 : n \geq 0\} \cup \\ & \{16\cdot3^n+1 : n \geq 0\}. \end{aligned}$$

By the above computations, the number of any of the forms $3^{n+1}+1$, $2\cdot 3^n+1$, $4\cdot 3^n+1$, or $8\cdot 3^n+1$ occurring in $A_{5\delta(2)}\setminus A_{4\delta(2)}$ is $5=4\cdot 3^0+1$. Lemma 5.5 shows that $\|16\cdot 3^n+1\|=9+3n$, and hence $\delta(16\cdot 3^n+1)=9-3\log_3(16+3^{-n})$, which allows us to check that none of these lie in $A_{6\delta(2)}$. Finally, checking the complexity of $32\cdot 3^k$ can be done with Lemma 5.4, and checking the complexity of $5\cdot 3^k$ can be done with Lemma 5.5. $\square$

To make later computations easier, let us observe here that $11\delta(2)<\delta(16\cdot 3^0+1)=\delta(17)<12\delta(2)$, and that for $n\geq 1$, $\delta(16\cdot 3^n+1)>12\delta(2)$.

In the above, for illustration, we explicitly considered and excluded numbers of the form $3\cdot 3^n+1$, $2\cdot 3^n+1$, etc., for large n, despite having already computed their complexities earlier. Henceforth, to save space, we will simply not consider a number if we have already computed its defect and seen it to be too high. E.g., in the above proof, we would have simply said, "By the main theorem and the above computations, $B_{6\delta(2)}\setminus B_{5\delta(2)}\subseteq\{32,5\}\cup\{8\cdot 3^n+1:n\geq 0\}$".

Proposition 5.11.

$$B_{7\delta(2)}=B_{6\delta(2)}\cup\{64,7,10\},$$

and the elements of $A_{7\delta(2)}$ have the complexities listed in Theorem 5.1.

Proof. By the main theorem and the above computations,

$$B_{7\delta(2)}\setminus B_{6\delta(2)}\subseteq\{64,7,10\}\cup\{32\cdot 3^n+1:n\geq 0\}\cup\{5\cdot 3^n+1:n\geq 0\}.$$

Lemma 5.5 shows that $\|32\cdot 3^n+1\|=11+3n$ and, for $n\geq 2$, $\|5\cdot 3^n+1\|=6+3n$. Hence $\delta(32\cdot 3^n+1)=11-3\log_3(32+3^{-n})$, and, for $n\geq 2$, $\delta(5\cdot 3^n+1)=6-3\log_3(5+3^{-n})$ which allows us to check that none of these lie in $A_{7\delta(2)}$. Finally, checking the complexities of $64\cdot 3^k$, $7\cdot 3^k$, and $10\cdot 3^k$ can be done via Lemma 5.4 (for 64 and 10) and Lemma 5.5 (for 7 and 10). $\square$

To make later computations easier, let us observe here that $\delta(32\cdot 3^n+1)>12\delta(2)$ for all n, and that for $n\geq 2$, $\delta(5\cdot 3^n+1)>12\delta(2)$ as well. Indeed, as we will see, from this point on, no new examples of multiplying by a power of 3 and then adding 1 will ever have complexity less than $12\delta(2)$.

Proposition 5.12.

$$B_{8\delta(2)}=B_{7\delta(2)}\cup\{128,14,20\},$$

and the elements of $A_{8\delta(2)}$ have the complexities listed in Theorem 5.1.

Proof. By the main theorem and the above computations,

$$\begin{aligned} B_{8\delta(2)}\setminus B_{7\delta(2)} \subseteq & \{128,14,20\}\cup\{64\cdot 3^n+1:n\geq 0\}\cup \\ & \{7\cdot 3^n+1:n\geq 0\}\cup\{10\cdot 3^n+1:n\geq 0\}. \end{aligned}$$

Lemma 5.5 shows that $\|64\cdot 3^n+1\|=13+3n$, $\|10\cdot 3^n+1\|=8+3n$, and, for $n\neq 0,2$, $\|7\cdot 3^n+1\|=7+3n$. Using this to check their defects, we see that none of these lie in $A_{8\delta(2)}$, or even $A_{12\delta(2)}$. Finally, checking the complexities of $128\cdot 3^k$, $14\cdot 3^k$, and $20\cdot 3^k$ can be done with Lemma 5.4. $\square$

Proposition 5.13.

$$B_{9\delta(2)}=B_{8\delta(2)}\cup\{256,28,40,19\},$$

and the elements of $A_{9\delta(2)}$ have the complexities listed in Theorem 5.1.

Proof. By the main theorem and the above computations,

$$\begin{aligned} B_{9\delta(2)} \setminus B_{8\delta(2)} \quad &\subseteq \quad \{256, 28, 40, 19\} \cup \{128 \cdot 3^n + 1 : n \geq 0\} \cup \\ &\qquad \{14 \cdot 3^n + 1 : n \geq 0\} \cup \{20 \cdot 3^n + 1 : n \geq 0\}. \end{aligned}$$

Lemma 5.5 shows that $\|128 \cdot 3^n + 1\| = 15 + 3n$, and for $n \geq 1$, $\|14 \cdot 3^n + 1\| = 9 + 3n$ and $\|20 \cdot 3^n + 1\| = 10 + 3n$. Using this to check their defects, we see that none of these lie in $A_{8\delta(2)}$, or even $A_{12\delta(2)}$. Finally, checking the complexities of $256 \cdot 3^k$, $28 \cdot 3^k$, and $40 \cdot 3^k$, and $19 \cdot 3^k$ can be done via Lemma 5.4 (for 256, 28, and 40) and Lemma 5.5 (for 28 and 19). ☐

Proposition 5.14.

$$B_{10\delta(2)} = B_{9\delta(2)} \cup \{512, 13, 1, 56, 80, 55, 38\} \cup \{3 \cdot 3^n + 1 : n \geq 3\},$$

and the elements of $A_{10\delta(2)}$ have the complexities listed in Theorem 5.1.

Proof. By the main theorem and the above computations,

$$\begin{aligned} B_{10\delta(2)} \setminus B_{9\delta(2)} \quad &\subseteq \quad \{512, 13, 1, 56, 80, 55, 38\} \cup \{3 \cdot 3^n + 1 : n \geq 3\} \cup \\ &\qquad \{256 \cdot 3^n + 1 : n \geq 0\} \cup \{28 \cdot 3^n + 1 : n \geq 0\} \cup \\ &\qquad \{40 \cdot 3^n + 1 : n \geq 0\} \cup \{19 \cdot 3^n + 1 : n \geq 0\}. \end{aligned}$$

We know $\delta(1) = 1$. Lemma 5.5 shows that $\|256 \cdot 3^n + 1\| = 17 + 3n$, $\|28 \cdot 3^n + 1\| = 11 + 3n$, $\|40 \cdot 3^n + 1\| = 12 + 3n$, and for $n \geq 1$, $\|19 \cdot 3^n + 1\| = 10 + 3n$. Using this to check their defects, we see that none of these lie in $A_{10\delta(2)}$, or even $A_{12\delta(2)}$. Finally, checking the complexities of $512 \cdot 3^k$, $13 \cdot 3^k$, $56 \cdot 3^k$, $80 \cdot 3^k$, $55 \cdot 3^k$, $38 \cdot 3^k$, and $(3^{n+1} + 1)3^k$ can be done via Lemma 5.4 (for 512, 56, 80, and 38) and Lemma 5.5 (for 55 and $3^{n+1} + 1$). ☐

Proposition 5.15.

$$\begin{aligned} B_{11\delta(2)} \quad &= \quad B_{10\delta(2)} \cup \{1024, 26, 112, 37, 160, 110, 76\} \cup \\ &\qquad \{2(3 \cdot 3^n + 1) : n \geq 3\} \cup \{2 \cdot 3^n + 1 : n \geq 4\}, \end{aligned}$$

and the elements of $A_{11\delta(2)}$ have the complexities listed in Theorem 5.1.

Proof. By the main theorem and the above computations,

$$\begin{aligned} B_{11\delta(2)} \setminus B_{10\delta(2)} \quad &\subseteq \quad \{1024, 26, 112, 37, 160, 110, 76\} \cup \{2(3 \cdot 3^n + 1 : n \geq 3)\} \cup \\ &\qquad \{2 \cdot 3^n + 1 : n \geq 4\} \cup \{512 \cdot 3^n + 1 : n \geq 0\} \cup \\ &\qquad \{13 \cdot 3^n + 1 : n \geq 0\} \cup \{56 \cdot 3^n + 1 : n \geq 0\} \cup \\ &\qquad \{80 \cdot 3^n + 1 : n \geq 0\} \cup \{55 \cdot 3^n + 1 : n \geq 0\} \cup \\ &\qquad \{38 \cdot 3^n + 1 : n \geq 0\} \cup \{(3 \cdot 3^n + 1)3^m + 1 : n \geq 3, m \geq 0\} \end{aligned}$$

Lemma 5.5 shows that for $m \geq 3$, $\|(3^{m+1} + 1)3^n\| = 1 + 3(m + 1) + 3n$, and that for $n \geq 1$, $\|512 \cdot 3^n + 1\| = 19 + 3n$, $\|56 \cdot 3^n + 1\| = 13 + 3n$, $\|80 \cdot 3^n + 1\| = 14 + 3n$, $\|55 \cdot 3^n + 1\| = 13 + 3n$, $\|38 \cdot 3^n + 1\| = 12 + 3n$, and that for $n \geq 2$, $\|13 \cdot 3^n + 1\| = 9 + 3n$. Using this to check their defects, we see that none of these lie in $A_{11\delta(2)}$, or even $A_{12\delta(2)}$. Finally, checking the complexities of $1024 \cdot 3^k$, $26 \cdot 3^k$, $112 \cdot 3^k$, $37 \cdot 3^k$, $160 \cdot 3^k$, $110 \cdot 3^k$, $76 \cdot 3^k$, $2(3^{n+1} + 1)3^k$, and $(2 \cdot 3^n + 1)3^k$ can be done via Lemma 5.4 (for 1024, 26, 112, 38, 160, 110, 76, and $2(3^{n+1} + 1)$) and Lemma 5.5 (for 37 and $2 \cdot 3^n + 1$). ☐

Proposition 5.16.

$$\begin{aligned} B_{12\delta(2)} &= B_{11\delta(2)} \cup \{2048, 25, 52, 224, 74, 320, 17, 220, 152, 73\} \cup \\ &\quad \{4(3\cdot 3^n+1) : n \geq 3\} \cup \{2(2\cdot 3^n+1) : n \geq 4\} \cup \\ &\quad \{4\cdot 3^n + 1 : n \geq 3\} \end{aligned}$$

and the elements of $A_{12\delta(2)}$ have the complexities listed in Theorem 5.1.

Proof. By the main theorem and the above computations,

$$\begin{aligned} B_{12\delta(2)} \setminus B_{11\delta(2)} &\subseteq \{2048, 25, 52, 224, 74, 320, 220, 152, 73\} \cup \\ &\quad \{4(3\cdot 3^n+1) : n \geq 3\} \cup \{2(2\cdot 3^n+1) : n \geq 4\} \cup \\ &\quad \{4\cdot 3^n+1 : n \geq 3\} \cup \{1024\cdot 3^n+1 : n \geq 0\} \cup \\ &\quad \{26\cdot 3^n+1 : n \geq 0\} \cup \{112\cdot 3^n+1 : n \geq 0\} \cup \\ &\quad \{37\cdot 3^n+1 : n \geq 0\} \cup \{160\cdot 3^n+1 : n \geq 0\} \cup \\ &\quad \{110\cdot 3^n+1 : n \geq 0\} \cup \{76\cdot 3^n+1 : n \geq 0\} \cup \\ &\quad \{2(3\cdot 3^n+1)3^m+1 : n \geq 3, m \geq 0\} \cup \\ &\quad \{(2\cdot 3^n)3^m+1 : n \geq 4, m \geq 0\} \end{aligned}$$

Lemma 5.5 shows that for $m \geq 3$ and $n \geq 1$, $\|2(3^{m+1}+1)3^n\| = 3+3(m+1)+3n$, and that for $m \geq 4$ and $n \geq 1$, $\|(2\cdot 3^m+1)3^n\| = 3+3m+3n$, and that $\|1024\cdot 3^n+1\| = 21+3n$, $\|112\cdot 3^n+1\| = 15+3n$, $\|160\cdot 3^n+1\| = 16+3n$, $\|76\cdot 3^n+1\| = 14+3n$, and that for $n \geq 1$, $\|26\cdot 3^n+1\| = 11+3n$, $\|110\cdot 3^n+1\| = 15+3n$, and that for $n \geq 2$, $\|37\cdot 3^n+1\| = 12+3n$. Using this to check their defects, we see that none of these lie in $A_{12\delta(2)}$. Finally, checking the complexities of $2048\cdot 3^k$, $25\cdot 3^k$, $52\cdot 3^k$, $224\cdot 3^k$, $74\cdot 3^k$, $320\cdot 3^k$, $220\cdot 3^k$, $152\cdot 3^k$, $73\cdot 3^k$, $4(3^{n+1}+1)3^k$, $2(2\cdot 3^n+1)3^k$, and $(4\cdot 3^n+1)3^k$ can be done via Lemma 5.4 (for 2048, 25, 52, 224, 74, 320, 220, 152, $4(3^{n+1}+1)$, and $2(2\cdot 3^n+1)$) and Lemma 5.5 (for 25, 73, and $4\cdot 3^n+1$). ☐

Combining all these propositions establishes Theorem 5.1.

6. Applications

We now make several applications of the classification obtained in Section 5. These are: (i) Classification of all integers n having defect $0 \leq \delta(n) \leq 1$ and finiteness of B_r for all $r < 1$; (ii) Determination of complexities $\|2^a\cdot 3^k\|$ for $a \leq 21$ and all k; (iii) Upper bounds on the number of integers $n \leq x$ having complexity $\delta(n) < r$, for any fixed $r > 0$.

6.1. Classifying the integers of Defect at most 1.

Using Theorem 5.1 we can classify all the numbers with defect less than 1, as follows:

Theorem 6.1. *The natural numbers n satisfying $\delta(n) < 1$ are precisely those that can be written in one of the following forms, and have the following complexities:*

1. 3^k *for* $k \geq 1$, *of complexity* $3k$
2. 2^a3^k *for* $a \leq 9$, *of complexity* $2a+3k$ *(for* a, k *not both zero)*
3. $5\cdot 2^a3^k$ *for* $a \leq 3$, *of complexity* $5+2a+3k$
4. $7\cdot 2^a3^k$ *for* $a \leq 2$, *of complexity* $6+2a+3k$
5. $19\cdot 3^k$ *of complexity* $9+3k$
6. $13\cdot 3^k$ *of complexity* $8+3k$
7. $(3^n+1)3^k$ *of complexity* $1+3n+3k$ *(for* $n \neq 0$*)*

Furthermore $n = 1$ *is the only number having defect exactly* 1.

Proof. This list includes all numbers in $A_{9\delta(2)}$, and some numbers in $A_{10\delta(2)}$. These in turn are determined by the corresponding lists for $B_{9\delta(2)}, B_{10\delta(2)}$, in the latter case (Proposition 5.14) checking the complexities to exclude the leaders $\{56, 80, 55, 38\}$. $\square$

Using this list one may deduce the following important fact.

Theorem 6.2. *For every* $0 < \alpha < 1$*, the set of leaders* B_α *is a finite set. For every* $\alpha \geq 1$*, the set* B_α *is an infinite set.*

Proof. The first part follows from the fact that each of the categories above has a finite set of leaders, and that the final list (7) has a finite number of sublists with defect smaller than $1 - \epsilon$, for any epsilon. The defects

$$\delta((3^n + 1)3^k) = (3n + 1) - 3\log_3(3^n + 1) = 1 - 3\log_3(1 + \frac{1}{3^n})$$

approach 1 from below as n approaches infinity. This also establishes that B_1 is an infinite set, giving the second part. $\square$

6.2. **The complexity of $2^a \cdot 3^b$ for small a.** The determination of A_r in Theorem 5.1 allows us to put lower bounds on the complexities of any numbers not in it. Thus for instance we have the following result.

Lemma 6.3. *Let* n *be a natural number and suppose that there is no* k *such that* $2^{n+9}3^k \in A_{n\delta(2)}$*. Then for any* $m \leq n+9$ *and any* k *(with* m *and* k *not both zero),* $\|2^m 3^k\| = 2m + 3k$.

Proof. It suffices to show that $\|2^{n+9}3^k\| > 2n + 3k + 17$, but by assumption,

$$\|2^{n+9}3^k\| > (n + 9)3\log_3 2 + 3k + n\delta(2) = 2n + 3k + 27\log_3 2 > 2n + 3k + 17,$$

and we are done. $\square$

This lemma immediately establishes Conjecture 1.3 for $a \leq 21$.

Proof of Theorem 1.7. From our classification, it is straightforward to check that $2^{21}3^k$ does not lie in $A_{12\delta(2)}$ for any k, so we can conclude for $m \leq 21$ and any k, with m and k not both zero, $\|2^m 3^k\| = 2m + 3k$. $\square$

6.3. **Counting the integers below x having defect at most r.** In our computations in Section 5, we used a small step size $\alpha = \delta(2)$, and kept our superset of A_r small by using a pruning step. In what follows, we will use a different trick to keep our supersets of A_r from getting too large. Instead of pruning, we will use step sizes arbitrarily close to 1.

Proposition 6.4. *Given any* $0 < \alpha < 1$*, and any* $k \geq 1$*, we have that* $B_{k\alpha}(x) = O_{k\alpha}((\log x)^{k-1})$*, and* $A_{k\alpha}(x) = O_{k\alpha}((\log x)^k)$.

Proof. We induct on k. Suppose $k = 1$; by Corollary 6.2, then $B_{k\alpha} = B_\alpha$ is a finite set, so $B_{k\alpha}(x) = O_{k\alpha}(1)$. Also, for any r, $A_r(x) \leq B_r(x)(\log_3 x)$; in particular, $A_{k\alpha}(x) = O_{k\alpha}(\log x)$.

So suppose it is true for k and we want to prove it for $k + 1$; we apply Proposition 4.4 with step size α. For convenience, let S_r denote the set of solid numbers b satisfying $\|b\| < r + 3\log_3 2$, as mentioned in the discussion after Theorem 4.4; for any r, this is a finite set.

In the case $k+1=2$,

$$\begin{aligned}B_{2\alpha}(x) &\leq B_\alpha(x)^3 + (A_\alpha(x)|S_{2\alpha}| + |T_\alpha|)(|B_\alpha|+1)\\ &= O_\alpha(1)^3 + O_\alpha(\log x) + O_\alpha(1)\\ &= O_{(k+1)\alpha}(\log x).\end{aligned}$$

In the case $k+1>2$,

$$\begin{aligned}B_{(k+1)\alpha}(x) &\leq \sum_{\substack{i+j=k+2\\ i,j\geq 2}} B_{i\alpha}(x)B_{j\alpha}(x) + (A_{k\alpha}(x)|S_{(k+1)\alpha}| + |T_\alpha|)(|B_\alpha|+1)\\ &= \sum_{\substack{i+j=k+2\\ i,j\geq 2}} O_{i\alpha}((\log x)^{i-1})O_{j\alpha}((\log x)^{j-1}) + O_{k\alpha}((\log x)^k) + O_\alpha(1)\\ &= O_{k\alpha}((\log x)^k).\end{aligned}$$

In either case, we also have $A_{(k+1)\alpha}(x) = O_{(k+1)\alpha}((\log x)^{k+1})$. This completes the proof.

□

Using this result we conclude:

Theorem 6.5. *For any number $r > 0$, $B_r(x) = \Theta_r((\log x)^{\lfloor r\rfloor})$, and $A_r(x) = \Theta_r((\log x)^{\lfloor r\rfloor+1})$.*

Proof. For the upper bound, it suffices to note that $r = (\lfloor r\rfloor+1)\frac{r}{\lfloor r\rfloor+1}$, and that $\frac{r}{\lfloor r\rfloor+1} < 1$, and apply Proposition 6.4.

For the lower bound, let $k = \lfloor r\rfloor$, and consider numbers of the form

$$N = ((\ldots((3\cdot 3^{n_k}+1)3^{n_{k-1}}+1)\ldots)3^{n_1}+1)3^{n_0}.$$

Then

$$\|N\| \leq 3(n_0+\ldots+n_k)+k$$

and since $\log_3 N \geq n_0+\ldots+n_k$, this means $\delta(N)\leq k$. Furthermore, if $n_0=0$ and $n_1>0$ then N is not divisible by 3 and so is a leader. It is then easy to count that there are at least $\binom{\lfloor \log_3 x\rfloor}{k+1} \gtrsim \frac{1}{(k+1)!}(\log_3 x)^{k+1}$ such N less than a given x, and at least $\binom{\lfloor \log_3 x\rfloor}{k} \gtrsim \frac{1}{k!}(\log_3 x)^k$ if we insist that N be a leader. □

An immediate consequence of Theorem 6.5 is Theorem 1.8 in the introduction.

Proof of Theorem 1.8. The existence of numbers of arbitrarily large defect follows from the fact that the set of integers of defect $< r$ has density zero. □

This result is a long way from proving a bound of type $\|n\| \nsim 3\log_3 n$.

7. Acknowledgements

The authors are indebted to Jānis Iraids and Karlis Podnieks for supplying a wealth of numerical data. We thank Jeffrey Lagarias for looking over an early draft of this paper and elucidating just what it was we were doing, as well as for other help with editing, and to Mike Zieve and David Rohrlich for providing assistance with early drafts of the paper. We thank Paul Pollack and Mike Bennett for pointing out the paper [12]. Most of all we thank Juan Arias de Reyna for greatly clarifying much of our work, suggesting improved notation, shortening some proofs, and helping extensively with structuring and editing of this paper. We thank the reviewer for very helpful comments.

References

[1] Harry Altman, Integer Complexity: The Integer Defect, in preparation.

[2] Harry Altman, Integer Complexity and Well-Ordering, in preparation.

[3] Harry Altman, Computation of Numbers with Integer Complexity Close to the Lower Bound, in preparation.

[4] J. Arias de Reyna, Complejidad de los números naturales, *Gaceta R. Soc. Mat. Esp.*, **3** (2000), 230–250.

[5] Richard K. Guy, Some suspiciously simple sequences, *Amer. Math. Monthly*, **93** (1986), 186–190; and see **94** (1987), 965 & **96** (1989), 905.

[6] Richard K. Guy, *Unsolved Problems in Number Theory*, Third Edition, Springer-Verlag, New York, 2004, pp. 399–400.

[7] Jānis Iraids, Kaspars Balodis, Juris Čerņenoks, Mārtiņš Opmanis, Rihards Opmanis, Kārlis Podnieks. Integer Complexity: Experimental and Analytical results, arXiv:1203.6462, 2012

[8] Jeffrey C. Lagarias, On ternary expansions of powers of 2, *J. London Math. Soc.*, **79** (2009), 562–588. *MR* 2508867.

[9] K. Mahler & J. Popken, On a maximum problem in arithmetic (Dutch), *Nieuw Arch. Wiskunde*, (3) **1** (1953), 1–15; *MR* **14**, 852e.

[10] Daniel A. Rawsthorne, How many 1's are needed?, *Fibonacci Quart.*, **27** (1989), 14–17; *MR* **90b**:11008.

[11] Vivek V. Srinivas & B. R. Shankar, Integer Complexity: Breaking the $\Theta(n^2)$ barrier, *World Academy of Science*, **41** (2008), 690–691

[12] C. L. Stewart, On the Representation of an Integer in Two Different Bases, *J. Reine Angew. Math.*, **319** (1980), 63–72.

[13] Joshua Zelinsky, An Upper Bound on Integer Complexity, in preparation