Algebraic factors of $b^n - 1$ and $b^n + 1$ — more than you might expect

March 5, 2016

Sam Wagstaff
Purdue University CS, Math, CERIAS
West Lafayette, Indiana

I thank Mikhail Atallah, Paul M Kuliniewicz and Usman Latif for their help with this work.

Background

The Cunningham Project factors integers $b^n \pm 1$, where $2 \leq b \leq 12$ and $1 \leq n <$ a few hundred or 1200 for $b = 2$.

The factors of these numbers arise in many math problems and in cryptography.

- Period length of decimal fractions. $1/37 = 0.027027\ldots$ since $37 \mid 999$, not 99 or 9. But $1/7 = 0.142857\ldots$ since $7 \mid 999999$.

- (Odd) perfect numbers: 6, 28, 496.

- Amicable pairs of numbers: 220, 284.

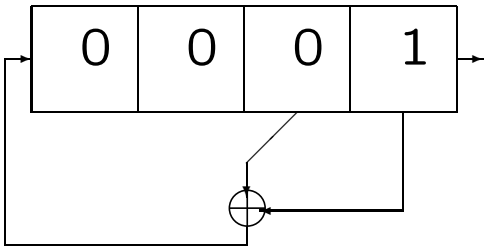
- Design of LFSRs.

- Constructing elliptic curves with nice properties. Add points. Compute pairings.

Linear Feedback Shift Registers

A *linear feedback shift register* is a device which generates a key stream for a stream cipher. It consists of an n-bit shift register and an XOR gate.

The inputs to the XOR gate are several bits selected (tapped) from fixed bit positions in the register. The output of the XOR gate is shifted in to the left side of the register.



Tap sequence $(t_1, t_2, t_3, t_4) = (0, 0, 1, 1)$.

Characteristic polynomial:
$1 + \sum_{i=1}^{4} t_i x^i = x^4 + x^3 + 1$.

The bits that are shifted out the right side of the register are the key stream. They are XORed with the plaintext to form the ciphertext (or vice versa to decrypt).

If there are n bits in the register, then you can choose positions to tap that make the bit stream have period $2^n - 1$, but you need to know the prime factors of $2^n - 1$ to do this.

Theorem. The period of an n-bit LFSR is maximal ($= 2^n - 1$) iff the characteristic polynomial is irreducible over $\mathbb{F}_2$, it divides $x^{2^n-1} + 1$, and it does not divide $x^d + 1$ for any proper divisor d of $2^n - 1$.

The RSA Public-Key Cipher

Choose two large primes p, q, so that $N = pq$ is too hard to factor.

Choose e with $\gcd(e, \phi(N)) = 1$, where $\phi(N) = (p-1)(q-1)$.

Via the extended Euclidean algorithm, compute d with $ed \equiv 1 \pmod{\phi(N)}$.

Discard p and q.

The public key is N, e. The private key is d.

Encipher M as $C = M^e \bmod N$. Decipher C as $M = C^d \bmod N$.

Sign M by $S = M^d \bmod N$. Verify the signature by $M = S^e \bmod N$.

The obvious attack on RSA is to factor N.

President Kennedy put locks on all nuclear weapons, with the President holding the keys.

The security of the (RSA) locks depends on the factoring problem being hard.

If you can factor large numbers, then you can arm nuclear weapons (or disarm them).

The cryptographic locks on the controls of the electrical grid use a similar system.

If you can factor large numbers, then you can turn off the lights in New York City.

Banks use RSA to validate wire transfers.

The problem of distinguishing prime numbers from composites, and of resolving composite numbers into their prime factors, is one of the most important and useful in arithmetic... The dignity of science seems to demand that every aid to the solution of such an elegant and celebrated problem be zealously cultivated.

— Carl Freidrich Gauss, *Disquisitiones Arithmeticae*, 1801.

Suppose, for example, that two 200-digit numbers p and q have been proved prime. . . . Suppose further that the cleaning lady gives p and q by mistake to the garbage collector, but that the product pq is saved. It must be felt as a defeat for mathematics that, in these circumstances, the most promising approaches [for finding p and q] are searching the garbage dump and applying mnemo-hypnotic techniques.

— Hendrik W, Lenstra, Jr. *Computational Methods in Number Theory*, 1981.

Table 10-

n	Prime Factors of 10^n - 1
1	3.3
2	3.3.11
3	3.3.3.37
4	3.3.11.101
5	3.3.41.271
6	3.3.3.7.11.13.37
7	3.3.239.4649
8	3.3.11.73.101.137
9	3.3.3.3.37.333667
10	3.3.11.41.271.9091
11	3.3.21649.513239
12	3.3.3.7.11.13.37.101.9901
13	3.3.53.79.265371653
14	3.3.11.239.4649.909091
15	3.3.3.31.37.41.271.2906161
16	3.3.11.17.73.101.137.5882353
17	3.3.2071723.5363222357
18	3.3.3.3.7.11.13.19.37.52579.333667
19	3.3.1111111111111111

Table 2-

n	Prime Factors of 2^n - 1
2	3
3	7
4	3.5
5	31
6	3.3.7
7	127
8	3.5.17
9	7.73
10	3.11.31
11	23.89
12	3.3.5.7.13
13	8191
14	3.43.127
15	7.31.151
16	3.5.17.257
17	131071
18	3.3.3.7.19.73
19	524287

Cyclotomic factors of $b^n - 1$

Let $\Phi_d(x) \in \mathbf{Z}[x]$ denote the dth *cyclotomic polynomial*. The degree of $\Phi_d(x)$ is $\phi(d)$.

$$\begin{aligned}
\Phi_1(x) &= x - 1 \\
\Phi_2(x) &= x + 1 \\
\Phi_3(x) &= x^2 + x + 1 \\
\Phi_4(x) &= x^2 + 1 \\
\Phi_p(x) &= x^{p-1} + x^{p-2} + \cdots + x + 1 \\
\Phi_6(x) &= x^2 - x + 1 \\
\Phi_8(x) &= x^4 + 1 \\
\Phi_9(x) &= x^6 + x^3 + 1 \\
\Phi_{10}(x) &= x^4 - x^3 + x^2 - x + 1 \\
\Phi_{12}(x) &= x^4 - x^2 + 1 \\
\Phi_{20}(x) &= x^8 - x^6 + x^4 - x^2 + 1.
\end{aligned}$$

$\Phi_d(x)$ is irreducible over $\mathbb{Q}$. Its zeros are the primitive d-th roots of 1, namely,
$\cos\frac{2\pi j}{d} + i \sin\frac{2\pi j}{d}$ with $\gcd(j, d) = 1$.

Let $n > 1$ be an integer.

We have $x^n - 1 = \prod_{d|n} \Phi_d(x)$ since every nth root of 1 is a primitive d-th root of 1 for some $d \mid n$.

Substituting $x = b$ into this formula gives a (partial) factorization of the integer $b^n - 1$.

Example: Since 9, 3, 1 are the divisors of 9,

$$x^9 - 1 = \Phi_9(x)\Phi_3(x)\Phi_1(x).$$

$$\begin{aligned} 6^9 - 1 &= \Phi_9(6)\Phi_3(6)\Phi_1(6) \\ &= (6^6 + 6^3 + 1)(6^2 + 6^1 + 1)(6 - 1) \\ &= (19 \cdot 2467)(43)(5). \end{aligned}$$

The factorization of $b^n - 1$ is nontrivial except when $b = 2$ and n is prime.

Cyclotomic factors of $b^n + 1$

Likewise, $x^n + 1$ can be partially factored as $\prod \Phi_d(x)$ where the product extends over integers d dividing $2n$ but not n (because $x^n + 1 = (x^{2n} - 1)/(x^n - 1)$).

Example: Since 12 and 4 divide 12, but not 6, we have

$$\begin{aligned} 5^6 + 1 &= \Phi_{12}(5)\Phi_4(5) \\ &= (5^4 - 5^2 + 1)(5^2 + 1) \\ &= (601)(2 \cdot 13) \end{aligned}$$

This gives a factorization of $b^n + 1$ which is nontrivial when n is not a power of 2.

We call the product formulas on this slide and the preceding one the *cyclotomic factorizations* of $b^n - 1$ and $b^n + 1$.

In the 1860s, Landry factored several numbers like $2^{50}+1$, $2^{54}+1$ and $2^{62}+1$ by trial division. He mentioned that the factorization

$$2^{58}+1 = 5 \cdot 107367629 \cdot 536903681$$

was much harder than the others.

In 1871, Aurifeuille rewrote this number as

$$2^{58}+1 = 536838145 \cdot 536903681,$$

where the first factor is $5 \cdot 107367629$. He observed that the two displayed factors are close together, their difference is 2^{16}, and the equation is the special case $k = 15$ of the identity

$$2^{4k-2}+1 = (2^{2k-1}-2^k+1)(2^{2k-1}+2^k+1).$$

Landry's task would have been much easier had he noticed this identity before he began to factor $2^{58}+1$ by trial division.

Aurifeuillian factorizations

Aurifeuillian factorizations are algebraic factorizations of $b^n \pm 1$ that go beyond the cyclotomic factorization of some of these numbers.

The simplest one is the identity

$$2^{4k-2} + 1 = (2^{2k-1} - 2^k + 1)(2^{2k-1} + 2^k + 1),$$

which factors the number $2^n + 1$, with $n = 4k - 2$, in a different way from the cyclotomic factorization of this number.

Example:

$$\begin{aligned} 2^{10} + 1 &= 2^{10} + 2^6 + 1 - 2^6 \\ &= (2^{10} + 2 \cdot 2^5 + 1) - 2^6 \\ &= (2^5 + 1)^2 - (2^3)^2 \\ &= (2^5 + 2^3 + 1)(2^5 - 2^3 + 1) \\ &= 41 \cdot 25 = (2, 10M) \cdot (2, 10L). \end{aligned}$$

Cyclotomic vs. Aurifeuillian factors

In summary, the cyclotomic factorization of $2^{10}+1 = 1025$ is

$$2^{10}+1 = \Phi_{20}(2)\Phi_4(2) = 205 \cdot 5$$

and the Aurifeuillian factorization is

$$2^{10}+1 = (2^5+2^3+1)(2^5-2^3+1) = 41 \cdot 25.$$

Aurifeuillian factors

The number $3^{3m}+1$ always has the cyclotomic factorization $(3^m+1)(3^{2m}-3^m+1)$, but the second factor could be prime. However, when $m=2k-1$ is odd, the second factor splits into two nearly equal parts:

$$3^{6k-3} + 1 = (3^{2k-1}+1) \times (3^{2k-1}-3^k+1)(3^{2k-1}+3^k+1).$$

In the Cunningham notation, the trinomial factor with -3^k is called "L" while the one with $+3^k$ is called "M".

Thus, "3,27L" is $3^9-3^5+1=19441$ and "3,27M" is $3^9+3^5+1=19927$. Note that 3,27L $\approx$ 3,27M. Also, $5^{5(2k-1)}-1 = (5^{2k-1}-1)\times L \times M$, where L, $M =$

$$5^{2(2k-1)}+3\cdot 5^{2k-1}+1 \mp 5^k(5^{2k-1}+1).$$

Table 5-

n	Prime Factors of 5^n - 1
1	2.2
3	(1) 31
5	(1) L.M
	L 11
	M 71
7	(1) 19531
9	(1,3) 19.829
11	(1) 12207031
13	(1) 305175781
15	(1,3) L.M
	L (5M) 181
	M (5L) 1741
17	(1) 409.466344409
19	(1) 191.6271.3981071
21	(1,3,7) 379.519499
23	(1) 8971.332207361361
25	(1,5L,5M) L.M
	L 9384251
	M 101.251.401

Applications of Aurifeuillian factors

A *primitive factor* of $b^n - 1$ is one which does not divide $b^i - 1$ for any $i < n$. Ditto for $b^n + 1$.

When Aurifeuillian factorizations occur, they cut across the cyclotomic factorizations and may be used to show that certain $b^n \pm 1$ have at least two primitive prime factors.

For example, $3^{27} + 1$ has the two primitive prime factors 19441 and 19927 (of $\Phi_{27}(3)$).

In fact, $3^{6k-3} + 1 = 3^{3(2k-1)} + 1$ always has at least two primitive prime factors when $k > 1$. That is, $\Phi_{6k-3}(3)$ is always composite.

A corollary is that there are infinitely many composite pseudoprimes n to any base $b > 1$, that is, $b^{n-1} \equiv 1 \bmod n$.

Another important application of Aurifeuillian factorizations is to help factor Cunningham-type numbers $b^n \pm 1$.

When b is a odd prime, the Aurifeuillian factorization is for $b^{b(2k-1)} - 1$ when $b \equiv 1 \bmod 4$ and for $b^{b(2k-1)} + 1$ when $b \equiv 3 \bmod 4$.

In 1996, Wagstaff factored

$$\begin{aligned} N &= (173^{173} - 1)/(173 - 1) \\ &= 347 \cdot 685081 \cdot P18 \cdot P176 \cdot P184, \end{aligned}$$

where $Pxxx$ denotes a prime with xxx decimal digits.

If one began naively to factor N, it would be easy to discover the three small prime factors, but no algorithm known at this time could split the product of the two large prime factors. However, N has an Aurifeuillian factorization that breaks it into two nearly equal pieces, each of which is easy to factor.

Gauss knew some Aurifeuillian factorizations.

Aurifeuille (1870s) proved Aurifeuillian factorizations exist for infinitely many b.
In 1962, Schinzel proved the existence of Aurifeuillian factorizations for $a^n - b^n$ in many cases, including some new cases for $b^n \pm 1$ that Aurifeuille missed. Because of their beauty and usefulness, some mathematicians have sought other algebraic identities similar to those Schinzel found. Some published "new" Aurifeuillian factorizations.

In 2007, Granville and Pleasants showed that Schinzel found all such identities, provided one accepts their (reasonable) definition of "such identities." However, they noted that some numerical examples from the Cunningham tables suggest that other such identities might exist.

For example,

$$\begin{aligned} & 6^{106}+1 = 37 \times 26713 \times \\ \times \quad & 1754369260046476588102446137364791189 17 \\ \times \quad & 1757871574183058771734553557555468706 41 \end{aligned}$$

with two nearly equal 39-digit factors (each close to $6^{49.148}$).

This example does not come from either cyclotomic or Aurifeuillian factorizations.

In this talk, we describe a systematic empirical search of the Cunningham tables for new algebraic identities beyond those discovered by Schinzel.

How the Cunningham Tables Were Searched

It is easy to discover the cyclotomic factorizations in a table of numbers $b^n \pm 1$ in which each number is factored. Look for repeated primes.

Aurifeuillian factorizations are slightly harder to find by inspecting tables of factored numbers. It helps if the factors of $b^n \pm 1$ are grouped to form nearly equal products. Here is an example.

We have

$$\begin{aligned} N &= 5^{35} - 1 \\ &= 2 \cdot 2 \cdot 11 \cdot 71 \cdot 211 \cdot 631 \cdot 4201 \cdot 19531 \cdot 85280581 \\ &= (2 \cdot 2 \cdot 19531)\,(71 \cdot 85280581)\,(11 \cdot 211 \cdot 631 \cdot 4201) \\ &= 78124 \cdot 6054921251 \cdot 6152578751 \\ &= (5^7 - 1) \cdot 6054921251 \cdot 6152578751. \end{aligned}$$

Note that $78124^2 = 6103359376$, so that the factors in the last line are close to $N^{1/5}$, $N^{2/5}$, $N^{2/5}$, respectively. The two 10-digit factors in the last line give the Aurifeuillian factorization.

We will seek new algebraic identities by grouping primes to form nearly equal products.

Suppose p_1, ..., p_k are all the [primitive] prime factors of $b^n - 1$ or of $b^n + 1$ for particular b and n.

We seek subsets $T \subset \{1, \ldots, k\}$ so that

$$\prod_{i \in T} p_i \approx \prod_{i \notin T} p_i.$$

In the example of $5^{35} - 1$, the algebraic factor $5^7 - 1$ would have to be have to be removed first. Then we would partition the set of primes $\{11, 71, 211, 631, 4201, 85280581\}$ into two subsets with nearly equal products.

Later we will explain a way to do this which does not require noticing that $5^7 - 1$ must be removed.

The first observation is that one need not do much arithmetic with multiprecise integers.

It is easier to take logs of all the primes, say, $a_i = \log p_i$, and convert the problem to this one involving only sums of real numbers.

Given positive real numbers a_1, ..., a_k, find subsets $T \subset \{1, \ldots, k\}$ so that

$$\sum_{i \in T} a_i \approx \sum_{i \notin T} a_i.$$

Some cases are easy. If the largest prime p_k is greater than the product of the other prime factors, then there is at most one useful subset T.

If the number k of prime factors (or addends in the equivalent real number problem) is small, then we can examine all 2^{k-1} subsets T containing 1.

But in the interesting part of the Cunningham tables, k may be more than 40, too large for a brute force search.

For example, $3^{1155} + 1$ has 44 prime factors, which may be seen at `factordb.com`.

We will actually solve a slightly more general problem. Let $c = \sum_{i=1}^{k} a_i$. Let α be a real number between 0 and 1.

We will find subsets $T \subset \{1, \ldots, k\}$ so that $\sum_{i \in T} a_i \approx \alpha c$. The original problem had $\alpha = 1/2$.

If we set $\alpha = 1/3$, then we will be seeking a partition of the prime factors of $b^n \pm 1$ into two sets so that the product of the primes in one set is approximately equal to the square of the product of the primes in the other set.

One might choose $\alpha = 2/5$ to discover the Aurifeuillian factorization of $5^{35}-1$ in the example above. (Using this α, one need not remove the factor $5^7 - 1$.)

With no loss of generality we may assume $0 < \alpha \leq 1/2$. In the known Aurifeuillian factorizations of $b^n \pm 1$, the $\alpha = a/b$, where $0 < a \leq b/2$ is an integer. We decided to examine these values of α and a few others.

To handle large k, we replaced the real numbers by small positive integers.

Choose a positive integer K_1 of convenient size.

Define integers $b_1, \ldots, b_k$ by $b_i = \lfloor a_i K_1/c + 0.5 \rfloor$, that is, b_i is $a_i K_1/c$ rounded off to the nearest integer.

Since $0 < a_i \leq c$ we have $0 \leq b_i \leq K_1$. Let $K = \sum_{i=1}^{k} b_i$.

Then

$$\begin{aligned} K &= \sum_{i=1}^{k} b_i \approx \sum_{i=1}^{k} a_i K_1/c \\ &= \frac{K_1}{c} \sum_{i=1}^{k} a_i = \frac{K_1}{c} c = K_1. \end{aligned}$$

Also,

$$\sum_{i \in T} b_i \approx \alpha K \quad \Leftrightarrow \quad \sum_{i \in T} a_i \approx \alpha c.$$

We handled the approximate equalities as percentage errors. For example, allowing a 1% error means

$$0.99\alpha c \leq \sum_{i \in T} a_i \leq 1.01\alpha c.$$

The problem for $b_1, \ldots, b_k$ with $\alpha = 1/2$ is called the *partition problem* in computer science. The general problem with arbitrary α is known as the *subset sum problem*. Wikipedia has articles on both problems. The algorithms book of Cormen, Leiserson and Rivest discusses the subset sum problem. Both problems are NP-complete, but in many cases heuristics solve the problems quickly.

We decided to use a variation of the polynomial time approximation algorithm in the Wikipedia article on the subset sum problem.

Our algorithm uses a $k \times (K + 1)$ matrix $X_{r,j}$ of linked lists of pairs (i, b) of integers with $i + b = j$.

The pairs contain information needed to recursively express j as a sum of input numbers b.

The pair $(0, b)$ in $X_{0,b}$ represents the number b in the input.

When $r > 0$, the pair (i, b) in $X_{r,(i+b)}$ represents the sum of the input number b and the number i whose representation as a sum of input numbers is given in the entry $X_{(r-1),i}$ in the previous row.

The next slide shows the algorithm we used.

for $r = 1$ to k { insert $(0, b_r)$ onto X_{1,b_r} }
for $r = 2$ to k {
  for $j = 0$ to K {
    if $X_{(r-1),j} \neq \emptyset$ {
      b_n = least b of $(\cdot, b)$ on $X_{(r-1),j}$
      for $\ell = n + 1$ to $k - 1$
      { insert (j, b_ℓ) onto $X_{r,(j+b_\ell)}$ }
    }
  }
}
for $r = 1$ to k {
  for $j \approx \alpha K$ {
    for each entry in $X_{r,j}$ {
      follow the pointers back to build T
      if $\sum_{i \in T} b_i \approx \alpha K$,
check whether $\sum_{i \in T} a_i \approx \alpha c$ and print T if so.
    }
  }
}

In the algorithm, the condition $j \approx \alpha K$ means that j is one of the 1, 2 or 3 integers within 1 of the real number αK, unless the percentage error is very large.

We may ignore the columns $X_{r,j}$ of the matrix with $j > \alpha K$ (or $j > \alpha K + 1$) because these columns do not matter since their ordered pairs represent sums greater than αK of input numbers.

Theorem The algorithm takes $O(k^2 K + S)$ steps, where S is the number of solutions to the approximate equality above.

Sometimes, as for $3^{1155}+1$, the algorithm runs for a long time because there are so many answers to the problem it solves. Perhaps one of the answers is an instance of a new algebraic factorization with infinitely many cases. This slowness cannot be avoided.

The Results

We searched the Cunningham tables for new algebraic factorizations.

A percentage error of 1% or 0.1%, the program gave too many approximate solutions. Therefore, we used 0.01%.

However, this choice gave too many solutions in the $2^n + 1$ table, so we sharpened the percentage error to 0.001% for this table.

We did the computations for the primitive parts only, for the full factorization of $b^n \pm 1$, and for the L and M parts of Aurifeuillian factorizations separately.

We also combined the sides of the Aurifeuillian factorizations and rediscovered them as a way of testing the program. It found all of them when n was large enough so that the difference $|\log L - \log M|$ was within the percentage error of $\log L$.

Isolated approximate splits are interesting curiosities, but our goal was to find new identities having infinitely many cases.

Hence we searched for partitions of the factors of $b^n \pm 1$ with fixed α, b and $\pm$ with exponents n in an arithmetic progression or the start of another infinite sequence.

We took special notice of identities in which the approximation was much better than the allowed percentage error.

One clue that an approximate split is new was that the factors from an earlier referenced line were not all on the same side of the partition. When all prime factors from each earlier referenced line were on the same side of the partition, there was usually an algebraic explanation.

Base 2

We tried $\alpha = 1/4$ and $1/2$. The base 2 tables are longer than those of other bases and contain many numbers with lots of factors. The numbers $2^{1155} \pm 1$ and $2^{2310} + 1$ each gave tens of thousands of approximate solutions.

Many numbers in the $2^n + 1$ table split in non-Aurifeuillian ways into two factors nearly as close as their Aurifeuillian factorizations.

For example, consider $2^{534} + 1$. The number 2,534L contains factors from 2,2M and 2,178M. The number 2,534M contains factors from 2,6M and 2,178L. Each of 2,534L, 2,534M also has three primitive factors. However, the product of 2,178L and the three primitive factors of 2,534L is almost as close to the product of 2,2M, 2,6M, 2,178M and the three primitive factors of 2,534M as are the two factors of the Aurifeuillian factorization of $2^{534} + 1$. Similar examples exist for $2^n + 1$ with $n = 582$, 594, 618, 678, 702, 714, 726, etc.

Base 3

We tried $\alpha = 1/4$, $1/3$ and $1/2$.

With $\alpha = 1/2$ there were good approximate splits for $3^n - 1$ for $n = 165$, 225, 315, 375, 405, 525, 555, 585 and 615. These numbers all have many prime factors and the splits appeared to be coincidences even though these n are all $\equiv 15 \pmod{30}$.

There are many ways to swap L and M factors between the Aurifeuillian factors of $3^{6k-3} + 1$ to produce good approximate splits different from the Aurifeuillian factorizations.

Base 10

We tried $\alpha = 1/10$, $1/5$, $3/10$, $2/5$, $1/2$ and $1/4$.

With $\alpha = 1/2$ there were many non-Aurifeuillian factorizations of $10^n + 1$ for $n = 330$, 390, 450, 510, 570, 630, 690 and 750. Note that these exponents form an arithmetic progression. Some of these factorizations are just Aurifeuillian factorizations with a few factors swapped between the L and M sides.

With $\alpha = 0.4$, we found cyclotomic factorizations similar to those for $7^{105} - 1$. Here is an example:

$$(10^{75}-1)^{0.4} \approx 10^{30} \approx \Phi_3(10)\Phi_{15}(10)\Phi_{25}(10).$$

There are other similar examples for $10^n - 1$ with n in the arithmetic progression 105, 135, 165, ... and for $10^n + 1$ with $n = 120$, 135, 180, 225, 300, 315 and 360. The other values of α that we tried did not give such examples in base 10.

Are these results just coincidences?

Any real number $0 < \alpha < 1$ may be approximated within 2^{-k} by a sum of distinct powers of 2: 2^{-i} with $0 \le i < k$ using the binary representation of α.

Likewise, given k positive real numbers a_1, ..., a_k, with $a_i \le a_{i+1} \le 2a_i$ for $1 \le i < k$, any positive real number less than their sum $c = \sum_{i=1}^{k} a_i$ may be approximated within a_1 by a sum of a_i with distinct i.

When the real numbers a_i are the logarithms of the prime factors of a typical integer N, they might satisfy the inequality $a_i \le a_{i+1} \le 2a_i$ for several i.

For a typical random N, the a_i will roughly form a geometric progression. Cunningham numbers have more than the typical number of prime factors due to the cyclotomic and Aurifeuillian factorizations.

If a Cunningham number $N = b^n \pm 1$ has j prime factors, then it will often be possible to find subsets of the set of prime factors whose product is within a factor of $1 + 2^{-j}$ of N^α for many $0 < \alpha < 1$.

Some Cunningham numbers have more than fifteen prime factors. For these numbers and for any $0 < \alpha < 1$, there are many partitions of the set of prime factors in which the product of the primes in one piece is within a factor of $1 + 2^{-15}$ of N^α. These partitions produced many splits within 0.01% that were coincidences.

Conclusion

The Cunningham Project has been in existence for more than a century. It now lists the factorizations of thousands of integers $b^n \pm 1$. It has spurred the development of many new factoring algorithms. Billions of hours of computer time have been devoted to the project. Hundreds of people have helped factor numbers $b^n \pm 1$.

But there has been surprisingly little analysis of the results. The present work is one of the few studies of the tables. In it we have "data-mined" the Cunningham tables for algebraic factorizations like those discovered by Aurifeuille and generalized by Schinzel.

This research has been a treasure hunt. As I studied the lists of approximate products, many times I thought I had found a new algebraic factorization. In all cases where there was any sort of regularity, it always turned out that the approximate equality could be explained by known cyclotomic or Aurifeuillian factorization formulas, perhaps with similar factors swapped between sides.

I now believe that Schinzel really did find all algebraic factorizations of the Cunningham numbers and that Granville and Pleasants adequately captured the notion of "such identities" and showed that there are no others.

We hope this work inspires others to analyze the Cunningham tables in different ways and discover new truths.