EUCLID'S THEOREM REDUX

BEN DELO AND FILIP SAIDAK

Abstract. We discuss properties of certain recursive sequences, constructed in ways that guarantee pairwise coprimality of their iterative factors. As a direct application, we obtain a couple of new proofs of Euclid's Theorem concerning the infinitude of the prime numbers. In particular, one class of proofs of the theorem can be obtained from the following estimate:

$$\omega\left(a^n-1\right) \geq \Omega(n),$$

which we establish for all $a \geq 2$ and $n \in \mathbb{N}$ with $\gcd(a-1, n) = 1$.

1. Introduction

Let $n \in \mathbb{N}$. Recall the definitions of the arithmetical functions $\omega(n)$ and $\Omega(n)$,

$$\omega(n) = \sum_{p|n} 1 \quad \text{and} \quad \Omega(n) = \sum_{p^\alpha || n} \alpha,$$

counting the number of *distinct* prime factors of n and the *total* number of prime factors of n, respectively. Also, recall that $\phi(n) = \#\{m \leq n : \gcd(m, n) = 1\}$ is the multiplicative Euler's totient function, defined in 1758 (see [8]). In what follows, p and q will denote prime numbers, and $\pi(x) = \sum_{p \leq x} 1$ will be the prime counting function. It was Euclid, in his epoch-making *Elements* [5], who noticed (see [5], Liber IX, Proposition 20) that for all $1 \leq i \leq k$ we have

$$\gcd\left(p_i, 1 + \prod_{i=1}^{k} p_i\right) = 1, \tag{$\star$}$$

which implies that the set of prime numbers $\{p_i\}$ cannot be finite; in other words: $\pi(x) \to \infty$ as $x \to \infty$. Over the course of the past two millennia, many proofs of this result have been discovered, see the original work of Goldbach [11], Euler [7], Kummer [16], Hermite [13], Stieltjes [25], Thue [26], and Erdős [4], or compendiums of related ideas [2], [21], and [19].

The property of coprimality, which plays a key role in Euclid's original argument, can be used in other ways to prove the theorem that today bears his name. One idea (suggested in [22]) makes use of two consecutive integers that are always coprime – and because n and $n+1$ are coprime for all $n \in \mathbb{N}$, we necessarily have $\omega(n(n+1)) \geq 1 + \omega(n)$. Iteratively repeating this construction, one immediately obtains a sequence of integers whose terms have ever-increasing numbers of distinct prime factors. This implies that there must exist infinitely many prime numbers. The main goal of this short paper is to generalize this simple idea and see how far one can extend it using purely elementary techniques.

2. Quadratic Iterations

In a way analogous to the above construction, let us start with the following observation: if x is an odd integer, then x and $x+2$ will always be coprime, so $x(x+2)$ will have more prime factors than each of x and $x+2$. Notice that $x(x+2) = x^2 + 2x = (x+1)^2 - 1$,

so iterating this construction, with that starting value $x = 1$, one generates the sequence: $1, 2^2 - 1, 2^4 - 1, 2^8 - 1, 2^{16} - 1, \ldots$ with at least one new prime factor in each iteration:

$$\begin{aligned}
2^{2^0} - 1 &= 1 \\
2^{2^1} - 1 &= 3 && = 3 \\
2^{2^2} - 1 &= 15 && = 3 \cdot 5 \\
2^{2^3} - 1 &= 255 && = 3 \cdot 5 \cdot 17 \\
2^{2^4} - 1 &= 65535 && = 3 \cdot 5 \cdot 17 \cdot 257 \\
2^{2^5} - 1 &= 4294967295 && = 3 \cdot 5 \cdot 17 \cdot 257 \cdot 65537 \\
2^{2^6} - 1 &= 18446744073709551615 && = 3 \cdot 5 \cdot 17 \cdot 257 \cdot 65537 \cdot (641 \cdot 6700417)
\end{aligned}$$

It follows that, for all $n \in \mathbb{N}$,

$$\omega(2^{2^n} - 1) \geq n. \tag{1}$$

This gives another proof of Euclid's Theorem.

NOTE 1: The integers $2^{2^n} - 1$ are a special kind of Mersenne numbers, studied at great length by Marin Mersenne during the 1630s and 1640s (see [18] for more detail), but have been considered before 300 BC, making their appearance already in Euclid's *Elements* ([5], Liber IX, Proposition 36) in connection with the even perfect numbers.

NOTE 2: Reversing the process behind the construction of (1) actually leads back to the famous proof of Euclid's Theorem via pairwise coprimality of the Fermat numbers $2^{2^n} + 1$ (first defined in [10]), given by Christian Goldbach [11] in 1730:

$$\begin{aligned}
2^{2^n} - 1 &= (2^{2^{n-1}} + 1)(2^{2^{n-1}} - 1) \\
&= (2^{2^{n-1}} + 1)(2^{2^{n-2}} + 1)(2^{2^{n-2}} - 1) \\
&= \cdots \\
&= (2^{2^{n-1}} + 1)(2^{2^{n-2}} + 1)(2^{2^{n-3}} + 1) \cdots (2^{2^0} + 1) \implies \\
2^{2^n} + 1 &= (2^{2^{n-1}} + 1)(2^{2^{n-2}} + 1)(2^{2^{n-3}} + 1) \cdots (2^{2^0} + 1) + 2.
\end{aligned}$$

Let us also remark that the identity $x^2 - y^2 = (x + y)(x - y)$ (also found in *Elements* [5], Liber II, Proposition 5), applied repeatedly in the above decomposition of the special type of the Mersenne numbers from (1), can be used in a more general way. For example, observe

$$3^{2^n} - 2^{2^n} = (3^{2^{n-1}} + 2^{2^{n-1}})(3^{2^{n-1}} - 2^{2^{n-1}}),$$

where the two factors are again relatively prime because any common divisor would have to divide their difference, which equals $2 \cdot 2^{2^{n-1}}$, and that is impossible. Also, note that there is nothing special about 2 and 3 as bases. As long as a and b are coprime (and of different parity), any common divisor of $a^{2^{n-1}} + b^{2^{n-1}}$ and $a^{2^{n-1}} - b^{2^{n-1}}$ would have to divide $2 \cdot b^{2^{n-1}}$, but because this divisor cannot be even, it would have to divide $b^{2^{n-1}}$, but then also $a^{2^{n-1}}$, contradicting the condition of the coprimality of a and b. Thus, splitting off coprime factors of $a^{2^m} - b^{2^m}$, with m running from n down to 0, one can deduce (see also Maji [17]):

$$\omega(a^{2^n} - b^{2^n}) \geq n, \tag{2}$$

whenever a and b are of opposite parity and satisfy $\gcd(a, b) = 1$.

3. Cubics and Beyond

On closer inspection, several extensions of (1) and (2) offer themselves. Let us look in more detail at a couple of those that appear most natural and easiest to analyze. The original quadratic iteration $x \to x(x+2)$, which created our coprime factors, is itself a special case of something more general: notice that if we repeatedly employ $x \to (x+1)^3 - 1$, then because $(x+1)^3 - 1 = x(x^2+3x+3)$, the iteration will produce a split into two factors x and x^2+3x+3, and these will be coprime as long as $\gcd(x,3) = 1$. In other words (renaming $x+1$ as a, for the sake of simplicity), if $\gcd(a-1,3) = 1$, then, just as in the quadratic case, we have

$$\omega(a^{3^n} - 1) \geq n. \tag{3}$$

The coprime factors of the double-exponentials in (3) can be listed explicitly (see Note 3 below), but first we look at a further extension of the idea to arbitrary powers. Let $s \in \mathbb{N}$. As above, it can be seen that the sth degree iteration $x \to (x+1)^s - 1$ will generate two factors x and $\Psi_s(x) := x^{s-1} + \binom{s}{s-1}x^{s-2} + \cdots + \binom{s}{2}x + \binom{s}{1}$, and these factors will be coprime as long as $\gcd(x,s) = 1$. Starting with $x = 1$, this immediately yields

$$\omega(2^{s^n} - 1) \geq n,$$

for all $s \geq 2$. More generally, for all $a, s \geq 2$, with $\gcd(a-1,s) = 1$, we get

$$\omega(a^{s^n} - 1) \geq n. \tag{4}$$

Note 3. We can explicitly exhibit at least n coprime factors of the integers of the form (4). This is especially useful in the special cases when $s = p$ for prime p. From Fermat's Little Theorem (see [9]), the condition $p \nmid (a-1)$ gives us: $1 \not\equiv a \equiv a^p \equiv a^{p^2} \equiv a^{p^3} \equiv \cdots \pmod{p}$, enabling us to write out the decomposition:

$$\begin{aligned} a^{p^n} - 1 &= (a^{p^{n-1}} - 1)\Psi_p(a^{p^{n-1}} - 1) = \\ &= (a^{p^{n-2}} - 1)\Psi_p(a^{p^{n-2}} - 1)\Psi_p(a^{p^{n-1}} - 1) \\ &= \cdots \\ &= (a-1)\prod_{i=1}^{n-1} \Psi_p(a^{p^i} - 1) = \prod_{i=0}^{n-1} \Psi_p(a^{p^i} - 1), \end{aligned}$$

where the values of the Ψ function in the last product are coprime by pairs.

Note 4. Considering the bound in (1), writing $x = 2^{2^n} - 1$, and taking natural logarithms twice, one gets nontrivial information about the growth of $\pi(x)$, namely:

$$\pi(x) > \frac{\log\log x}{\log 2}.$$

With a change of base to $a > 2$, the bound stays essentially unchanged. The same is true for more elaborate iterative constructions $x_{k+1} = f(x_k)$, analogous to those discussed above, except incorporating *groups* of coprime integers $< x_k$. For example, modifying $x \to x(x+2)$ to $x \to x\Theta(x)$, where $\Theta(x)$ is a product of integers $< x$ and coprime to x, if in the kth iteration one could generate b_k new coprime factors, then up to $a^G = (\cdots((a^{b_1})^{b_2})^{b_3}\cdots)^{b_k}$, we would be guaranteed existence of at least $A = b_1 + b_2 + \cdots + b_k$ different prime numbers; i.e.,

$$\pi\left(a^G\right) \geq A \geq k\sqrt[k]{G},$$

by the elementary AM-GM inequality (see [15]). Choosing $x = a^G$, this can be rewritten as $\pi(x) \gg k(\log x)^{1/k}$, the right side of which, with growing k, attains its minimum value when

$$0 = \left(ke^{\frac{\log\log x}{k}}\right)' = e^{\frac{\log\log x}{k}} - k\frac{\log\log x}{k^2}e^{\frac{\log\log x}{k}} = \left(1 - \frac{\log\log x}{k}\right)(\log x)^{1/k},$$

i.e., when $k = \log\log x$. Unfortunately, this means that we cannot get a bound better than $\pi(x) \gg \log\log x$. This seems to be a limitation of all constructions of the above type. Any substantially better estimates would require more detailed information concerning the behavior of the iterations of Euler's ϕ function and its generalizations.

4. General Result

Now, we combine some of the above ideas to prove our main result.

Theorem 4.1. *For $a \geq 2$ and $n \in \mathbb{N}$, with $\gcd(a-1, n) = 1$, we have:*

$$\omega\left(a^n - 1\right) \geq \Omega(n). \tag{5}$$

Proof. Write $n = 2^\alpha M$, where M is odd. We handle the even components first. Clearly,

$$\begin{aligned} a^n - 1 &= a^{2^\alpha M} - 1 = (a^{2^{\alpha-1}M} + 1)(a^{2^{\alpha-1}M} - 1) \\ &= (a^{2^{\alpha-1}M} + 1)(a^{2^{\alpha-2}M} + 1)(a^{2^{\alpha-2}M} - 1) \\ &= \cdots \\ &= (a^M - 1)\prod_{i=0}^{\alpha-1}(a^{2^i M} + 1), \end{aligned}$$

where every step of the decomposition produces coprime factors, because for $i \geq 1$ we have $\gcd(a^{2^i M} + 1, a^{2^i M} - 1) = \gcd(a^{2^i M} - 1, 2) = \gcd(a-1, 2) = 1$, by assumption. Therefore,

$$\omega\left(a^n - 1\right) \geq \omega(a^M - 1) + \alpha = \omega(a^M - 1) + \Omega(2^\alpha).$$

Now, to handle $a^M - 1$, we proceed by induction on the number of distinct prime factors of M, with (4) serving as the base step. Let us assume that the bound in (5) is true for all integers with exactly k prime factors. Then, any M with $k+1$ prime factors can be written as $M = mq^\Delta$, where q is its smallest prime factor, $\omega(m) = k$, and $\gcd(q, m) = 1$. Again, by Fermat's Little Theorem, for all $i \geq 0$, $a^{mq^i} - 1 \equiv a^m - 1 \not\equiv 0 \pmod{q}$, because if $q \mid a^m - 1$, then $q \mid (a^m - 1)^{q^\Delta}$, and (because of the Binomial Theorem) q would divide $a^{mq^\Delta} - 1 = a^M - 1$, which would mean $(q-1) \mid M$, by Euler's Theorem [8]. This is impossible, because q is the smallest prime factor of M. Thus, the decomposition outlined in Note 4 can be applied:

$$\begin{aligned} a^M - 1 &= a^{mq^\Delta} - 1 = (a^{mq^{\Delta-1}} - 1)\Psi_q(a^{mq^{\Delta-1}} - 1) \\ &= (a^{mq^{\Delta-2}} - 1)\Psi_q(a^{mq^{\Delta-2}} - 1)\Psi_q(a^{mq^{\Delta-1}} - 1) \\ &= \cdots \\ &= (a^m - 1)\prod_{i=0}^{\Delta-1}\Psi_q(a^{mq^i} - 1), \end{aligned}$$

where, once again, all the factors are coprime by pairs. But this implies

$$\omega(a^M - 1) \geq \Omega(a^m - 1) + \Delta = \Omega(M),$$

by the inductive assumption concerning m. Hence,

$$\omega\left(a^n - 1\right) \geq \omega(a^M - 1) + \Omega(2^\alpha) \geq \Omega(M) + \Omega(2^\alpha) = \Omega(n),$$

as desired. This finishes the proof. □

NOTE 5. Maximal values of $\Omega(n)$, unlike those of $\omega(n)$, will grow to infinity, whether or not there exist infinitely many prime numbers. Therefore, for any $a \geq 2$, the bound in (5) implies

$$\max_{n \leq x} \omega(a^n - 1) \geq \max_{\substack{n \leq x \\ \gcd(a-1,n)=1}} \Omega(n) \to \infty$$

as $x \to \infty$. Equivalently, each sequence $\{a^n - 1\}_{n=1}^{\infty}$ confirms Euclid's Theorem.

NOTE 6. If $\gcd(a, b) = 1$, with positive integers a and b of opposite parity, then we expect the following general result to be true for all exponents $\alpha, \beta \in \mathbb{N}$:

$$\omega\left(a^\alpha - b^\beta\right) \geq \Omega(\gcd(\alpha, \beta)). \tag{6}$$

This generalizes (5). Due to some cumbersome congruence restrictions and the technical nature of analysis that is involved, we leave claim (6) as a conjecture. But, we remark that the Prime Number Theorem, $\pi(x) \sim x/\log x$ (see [14]), rewritten in the form $\sum_{p \leq x} \log p \sim x$, suggests that $\prod_{p \leq x} p \gtrsim e^x$. This means that any number with at least ω prime factors must be much larger than e^ω. Therefore, if one could establish (6), it would also follow that

$$a^\alpha - b^\beta > e^{\Omega(\gcd(\alpha,\beta))}, \tag{7}$$

a bound related to the Catalan and abc conjectures (see [20] and [1], respectively).

NOTE 7. If $\gcd(a, b) > 1$, then the powers a^α and b^β will have *extra* prime factors in common, and those will likely result in $\omega(a^\alpha - b^\beta)$ being even larger than when a^α and b^β are coprime. This is what one expects to happen in most, but not all cases. Heuristically, the same should be true of the bound in (5) of our main theorem, in situations when $\gcd(a - 1, n) > 1$. However, the iterative coprimality method we have employed in our proof of (5) will not work in these more general cases, because factoring out the common divisors in every step destroys the recursive process. So, in a way similar to the discussion in NOTE 4, although the bound in (5) is not optimal, its usefulness lies in its uniformity and simplicity.

5. Acknowledgements

We thank Numberphile [12] for inspiring us to revisit Euclid's Theorem, Caroline Kane for some helpful feedback, and the referee for several useful comments.

References

[1] J. Browkin, *The abc-conjecture*, Number Theory, 75–105, Trends Math., Basel, 2000.
[2] L. E. Dickson, *History of the Theory of Numbers*, Vol. 1, Chapter XVIII, 2nd ed., Chelsea Publishing, New York, 1992.
[3] G. Effinger and G. L. Mullen, *Two extended Euler functions with applications to latin squares and bases of finite field extensions*, Bulletin of the ICA, **85** (2019), 92–111.
[4] P. Erdős, *Über die Reihe* $\sum \frac{1}{p}$, Mathematica, Zutphen B, **7** (1938), 1–2.
[5] Euclid, *Elements*, Alexandria, c. 300 BC (see T. L. Heath, Thirteen Books of Euclid's Elements, Dover, New York, 1956).
[6] L. Euler, *De summis serierum reciprocarum* (1734), Comm. Acad. Sci. Petropolitanae **7** (1740), 123–134, (reprinted in Opera Omnia, Vol. 14, 73–86).
[8] L. Euler, *Theoremata arithmetica nova methodo demonstrata* (1758), Novi Comm. Acad. Sci. Petropolitanae, **8**, (1763) 74–104, (also in Opera Omnia, Vol. 2, 531–555).

[7] L. Euler, *Variae observationes circa series infinitas* (1737), Comm. Acad. Sci. Petropolitanae **9** (1744), 160–188, (reprinted in Opera Omnia, Vol. 14, 217–244).
[10] P. Fermat, *Letter to Frénicle*, August 1640, and *Letter to Pascal*, August 1654, Œuvres de Fermat, **2**, 206 & 309, Tannery & Henry (eds.), 1894.
[9] P. Fermat, *Letter to Mersenne*, June 1640, Œuvres de Fermat, **2**, 198, Tannery & Henry (eds.), 1894.
[11] C. Goldbach, *Letter of Euler*, July 31, 1730 (see Opera Omnia IV, A–4, Lemmermeyer & Mattmüller (eds.), Basel 2011).
[12] J. Grime, *Infinite Primes*, (Numberphile, YouTube), April 23, 2013.
[13] C. Hermite, *Note au sujet de la communication de M. Stieltjes ...*, Comptes Rendus Acad. Sci., **101** (1885), 112–115.
[14] G. J. O. Jameson, *The Prime Number Theorem*, LMS Student Texts 53, Cambridge University Press, Cambridge, 2003.
[15] P. P. Korovkin, *Inequalities*, Section 1.2, Mir Publishers, Moskva, 1975.
[16] E. E. Kummer, *Neuer elementarer Beweis des Satzes ...*, Monatsber. Preuß Wiss., (1878), 777–778.
[17] B. Maji, *A new proof of the infinitude of primes*, Resonance, **20**, no. 12 (2015), 1128–1135.
[18] M. Mersenne, *F. Marini Mersenni minimi Cogitata Physico Mathematica*, Praefatio Generalis No. 19, 1644.
[19] W. Narkiewicz, *The Development of Prime Number Theory*, Springer-Verlag, New York, 2000, 1–10.
[20] P. Ribenboim, *Catalan's Conjecture*, Academic Press, Boston, 1994.
[21] P. Ribenboim, *The New Book of Prime Number Records*, Springer-Verlag, New York, 1996.
[22] F. Saidak, *A new proof of Euclid's theorem*, Amer. Math. Monthly, **113**, no. 10 (2006), 937–938.
[23] F. Saidak, *A note on Euclid's theorem concerning the infinitude of the primes*, Acta Univ. M. Belii Ser. Math., **24** (2016), 59–60.
[24] V. Schemmel, *Über relative Primzahlen*, Journal für die Reine und Angew. Math., **70** (1869), 191–192.
[25] T. J. Stieltjes, Étude Bibliographique. sur la théorie des nombres, Annales Fac. Sci. de Toulouse, **4** (1890), 1–103.
[26] A. Thue, Mindre meddelelser II, Archiv for Math., Kristiania, **19**, no. 4 (1897), 1–5.

MSC2010: 11A05, 11A41, 11A51

Worcester College, Oxford, OX1 2HB, UK
Email address: ben.delo@worc.ox.ac.uk

Department of Mathematics, University of North Carolina, Greensboro, NC 27402, U.S.A.
Email address: f_saidak@uncg.edu