The factorisation of $2^{67} - 1$

PETER SHIU

1. *Introduction*

Frank Nelson Cole (1861-1921), a prominent American mathematician, was long time secretary of the American Mathematical Society (AMS) and editor-in-chief of its *Bulletin*. Upon his retirement two prizes, one in number theory and one in algebra, were established in his honour by the AMS. Cole is also remembered for a legendary achievement recorded by E. T. Bell in [1]:

> At the October, 1903 meeting in New York of the American Mathematical Society, Cole had a paper on the program with the modest title "On the factoring of large numbers". When the chairman called on him for his paper, Cole − who was already a man of very few words − walked to the board, and, saying nothing, proceeded to chalk up the arithmetic of raising 2 to the sixty-seven power. Then he carefully subtracted 1. Without a word he moved over to a clear space on the board and multiplied out, by longhand, 193707721 × 761838257287. The two calculations agreed. ... For the first and only time on record, an audience of American Mathematical Society vigorously applauded the author of a paper delivered before it. Cole took his seat without having uttered a word. Nobody asked him a question.

This is very likely a distorted story, mainly because Bell was prone to give Hollywood scenarios on what mathematicians actually do. Thus the veracity of the biographical sketches of mathematicians in his more widely read book [2] has been questioned by historians of mathematics, with many saying that he did not distinguish much between anecdote and history. In particular, the entries on Evariste Galois and Georg Cantor are said to be inaccurate, fanciful and even fictitious, and the item on Sofia Kovalevskaya has been condemned as an 'infuriatingly patronising, innuendo-laden mistreatment'. Indeed many now regard Bell as the legend maker of the history of mathematics, giving distorted impressions of the characters involved at every opportunity. Nevertheless, it has to be said that the book [2] did inspire some notable mathematicians, including Freeman Dyson, John Nash, Julia Robinson and Andrew Wiles, to take up a life of mathematics.

We consider how Cole arrived at the factorisation of $2^{67} - 1$ by studying his paper [3], which was 'Read before the American Mathematical Society, October 31, 1903', and published in December of the year. Much work, of course, has been done on the problem of factorisation of large numbers since the time of Cole, and readers who wish to know the history and development of the subject should read the very interesting article [4] by one of the leading experts.

https://doi.org/10.1017/mag.2014.13 Published online by Cambridge University Press

2. *Mersenne numbers*

The numbers $M_n = 2^n - 1$ have attracted much attention as regards their factorisation ever since Marin Mersenne gave some curious misinformation on them, with n being an odd prime $p \leqslant 257$, in the preface of his *Cogitata physico-mathematica* (1644). Some two hundred years later Edouard Lucas (1842-1891) made a serious study of Mersenne numbers. In particular, he devised an algorithm for testing the primality of M_p, and established in 1876 that M_{127} is indeed prime, as Mersenne claimed, and he also showed that, contrary to Mersenne's claim, M_{67} is actually composite, although his algorithm does not deliver the factors. Indeed the factors were first revealed by Cole during the AMS meeting in 1903 mentioned in the introduction. With a significant enhancement of Lucas's method by D. H. Lehmer in 1931, variants of the Lehmer-Lucas primality test are now fast algorithms for finding Mersenne primes and, in modern times, the largest known prime has almost always been the largest known Mersenne prime; moreover, the current ten largest known primes are Mersenne primes.

Finding factors of a composite Mersenne number M_p is much more difficult. Cole wrote in [3]: 'Lucas presented a brilliant method for ascertaining whether a given number $[M_p]$ is prime … based on this method, which however does not furnish the means of actually finding factors of the number examined.'

We shall make use of the following useful fact:

For an odd prime p, every prime factor of M_p is of the form $2kp + 1$ and $8\ell \pm 1$.

Readers who want to have a better understanding of the following should try to establish this fact, which makes use of Fermat's Little Theorem and the quadratic character of 2. We now try to (re)discover the least prime divisor P of the number

$$M = M_{67} = 2^{67} - 1 = 147573952589676412927, \quad M = PQ, \quad P < Q.$$

Calculations by Lucas already hinted that M is a product of two distinct primes, so that Q is also a prime, but we shall not assume this. From the useful fact above, we know that P satisfies the congruences $P \equiv 1 \pmod{134}$ and $p \equiv \pm 1 \pmod 8$, so that, by the Chinese remainder theorem, it has to lie in one of the two arithmetic progressions

$$P = 536k + h, \qquad k = 1, 2, \ldots, \qquad h = 1 \text{ or } 135; \tag{1}$$

we denote by Ω the finite set of primes P satisfying (1) and the bound $P < \sqrt{M} < 2^{67/2} < 12148002000$. With computing machines, perhaps a brute force search for the unique $P \in \Omega$ with $P \mid M$ may not be that formidable. However, back in 1900, arithmetic calculations could only be done by hand, so that Cole had to use mathematics to reduce the amount of computations to accomplish the task.

https://doi.org/10.1017/mag.2014.13 Published online by Cambridge University Press

3. *The methods of Fermat and Gauss*

The simplest method for the factorisation of n is by trial division, checking whether there is a prime divisor $p \leqslant \sqrt{n}$, and n can be declared prime if no such p exists. However, the method is useful only if the least prime divisor is small, and it becomes hopelessly inefficient when the divisor exceeds $\sqrt[3]{n}$, say. If it is known in advance that n has a divisor near $\sqrt{n}$, then it is much better to apply Fermat's method by trying to decompose n as a difference of two squares, say $A^2 - B^2$, which has the factorisation $(A + B)(A - B)$. For example, suppose that we have reason to believe that the prime divisors of 4886383 are not small. Then, instead of trial division, we may consider the smallest square which exceeds 4886383, namely $2211^2 = 4888521$ with the excess 2138, which is not a square; but the next square $2212^2 = 4892944$ has the excess 6561, which is 81^2. Thus

$$4886383 = 2212^2 - 81^2 = (2212 + 81)(2212 - 81) = 2293 \times 2131.$$

It is well explained in [4] that, following radical ideas from M. Kraitchik in the 1920s, and later from sophisticated theories in algebraic number theory and much else, Fermat's method has been much enhanced, and is now considered the basis of some very powerful methods of factorisation in our time.

Gauss made a serious study of the problem of factorisation, offering two methods which are based on the use of quadratic residues, and the law of quadratic reciprocity—this and much else were first established in 1801 in his *Disquisitiones Arithmeticae* [5]. As the author explained in §330 of this famous text, the foundation of his first method is the following useful observation:

> If a is a quadratic residue modulo n, then a is also a quadratic residue modulo every prime divisor of n.

Thus, on finding a quadratic residue a modulo n, we know that any prime divisor of n has to belong to the set of primes for which a is a quadratic residue. Moreover, from the theory of quadratic residues, such primes lie in half of the relevant arithmetic progressions with the modulus a or $4a$. Consequently once we have found a quadratic residue a modulo n, we should be able to eliminate about half of the elements from Ω the set of possible prime divisors of n.

4. *A first attempt at the factorisation of M*

Cole found that -3 is a quadratic residue modulo M, and we know that primes p for which -3 is a quadratic residue satisfy $p \equiv 1 \pmod 3$. We can therefore replace the search for P in (1) by the two thinner arithmetic progressions

$$P = 1608k + h, \qquad k = 1, 2, \ldots, \qquad h = 1 \text{ or } 1207; \qquad (2)$$

updating Ω with (2) instead of (1) we see that $|\Omega|$ is about half of what it was, but trial division is still a formidable task without machines. Let us

https://doi.org/10.1017/mag.2014.13 Published online by Cambridge University Press

quote Cole again:

'For the number $[M]$ I found the following apparently irreducible quadratic remainders:

2, −3, −7, 13, −23·53, 37, 41, 61, −67, −71, 23·83, 89, 97, 101, −23·113, −127, 137, 23·151, −23·157, 23·167, 173, 181, 23·191, −23·193, …

With the help of the small prime remainders, fortified by the presumption that −23 and therefore 53, −83, … were non-remainders, I had no difficulty in sifting the first sixteen million natural numbers, but the search proved without result.'

Cole did not mention the method by which these quadratic residues were found, but Gauss had already given three methods on how to find residues in §332 of [5]. The appearance of ±23 as a factor in several of the residues, together with the apparent failure to isolate either as a residue, suggest that M is composite, and that ±23 are non-residues (in fact $\left(\frac{\pm 23}{P}\right) = -1$ and $\left(\frac{\pm 23}{Q}\right) = -1$). The 'therefore' is a consequence of the product rule concerning residues and non-residues, and the 'sixteen million' corresponds to the halving of $|\Omega|$ when each of the 24 residues is considered, with $2^{24} = 2^4 2^{10} 2^{10} = 16 \times 1024 \times 1024 \approx 16 \times 10^6$. The hindsight of knowing that $P < 2 \times 10^8$ perhaps indicates that the employment of four more residues might have led to a successful determination of P. However, the work involved in an elimination process becomes more tedious as the size of the residue grows, and any newly found residues may no longer be small. Indeed, earlier in the paper Cole wrote that if the smallest factor exceeds 10^8 then 'a much more expeditious method is to employ the known quadratic remainders to reduce [the number] to the form $\left(\frac{u+v}{2}\right)^2 - \left(\frac{u-v}{2}\right)^2 = uv$, the factors u and v being so chosen that their difference $u - v$ is as small as possible.'

Thus, giving up on the idea of having more residues, Cole decided to change tack and 'the known' residues to Fermat's method instead.

5. *A successful factorisation of M*

In our notation, we have $(P, Q) = (v, u)$, so that $A = \frac{1}{2}(u + v)$ and $B = \frac{1}{2}(u - v)$ in Fermat's method, and

$$M = A^2 - B^2 = PQ, \qquad P = A + B, \qquad Q = A - B.$$

By Fermat's Little Theorem, $M = 2^{67} - 1 \equiv 2 - 1 \equiv 1 \pmod{67}$, so that $P \equiv Q \equiv 1 \pmod{67}$, and hence 67^2 divides

$$(P - 1)(Q - 1) = PQ + 1 - (P + Q) = 2^{67} - 2A.$$

Therefore $A \equiv 2^{66} \equiv 671 \pmod{67^2}$. We also find that $8 \mid A$, and that if $A \equiv 1 \pmod 3$ then $A \equiv 10, 37, 46, 64 \pmod{81}$. A similar argument shows that $A \equiv \pm 1 \pmod 5$, $A \equiv \pm 1, \pm 3 \pmod 7$, $A \equiv 0, \pm 1, \pm 3, \pm 6 \pmod{13}$. It then follows that $A \equiv r \pmod D$, where

$$D = 67^2 \cdot 8 \cdot 5 \cdot 7 \cdot 13 \cdot 81 = 1323536760,$$

https://doi.org/10.1017/mag.2014.13 Published online by Cambridge University Press

and r is one of the values obtained from the Chinese remainder theorem applied to the various systems of simultaneous congruences corresponding to all the possible cases. Cole found that the system

$$A \equiv -1 \pmod 5, \quad A \equiv 3 \pmod 7, \quad A \equiv 1 \pmod{13}, \quad A \equiv 46 \pmod{81},$$

which corresponds to $r = 1160932384$, delivers the solution with $A = xD + r$ when $x = 287$. One can use the square test to find the value for x, but Cole returned to the process of employing his residues for elimination, and his paper then ends in a flurry with

'Sifting with 23, 37, 41, 53, 61, I find that $x = 287$ remaining. A few further tests suffice to show that the goal is near. And, in fact,

$$287 \cdot 1323536760 + 1160932384 = 381015982504,$$

and

$$\begin{aligned} 2^{67-1} &= 381015982504^2 - 380822274783^2 \\ &= 193707721 \times 761838257287.' \end{aligned}$$

6. *Three years of Sundays*

While trying to understand how Cole was led to his system of congruences involving A, we realised that, instead of sifting with quadratic residues, he could also have just applied an enhancement of Fermat's method together with the square test. The realisation led to the content of [6].

It is said that, when asked how long did the factorisation took, Cole's reply was "Three years of Sundays". We have to admire the man for his effort and tenacity.

References

1. E. T. Bell, *Mathematics; queen and servant of sciences*, New York, McGraw-Hill (1951).
2. E. T. Bell, *Men of Mathematics* New York, Simon and Schuster (1937).
3. F. N. Cole, On the factoring of large numbers, *Bull. Amer. Math. Soc.*, **10** (3) (1903) pp. 134-137.
4. Carl Pomerance, A tale of two sieves, *Notices of the Amer. Math. Soc.* **43** 12 (1996) pp. 1473-1485.
5. C. F. Gauss, *Disquisitiones Arithmeticae* (1801), translated by Arthur A. Clarke, Yale University Press (1966).
6. Peter Shiu, Fermat's method of factorisation, *Math. Gaz.* **99** (March 2015) pp. 97-103.

doi: 10.1017/mag.2014.13

PETER SHIU
353 Fulwood Road, Sheffield S10 3BQ
e-mail: *p.shiu@yahoo.co.uk*

https://doi.org/10.1017/mag.2014.13 Published online by Cambridge University Press