Exploring Polynomial Identities to Factor Mersenne Numbers

Scott Stetson

October 27th 2022

Mersenne numbers are numbers that are of the form $2^n - 1$ for some $n \in \mathbb{N}$. They were first studied by Euclid in their connection with perfect numbers. Although Euclid wrote Mersenne numbers as the sum $1 + 2 + 2^2 + \cdots + 2^{n-1}$ which is equal to $2^n - 1$ by the identity below

$$(x-1)(x^{n-1} + x^{n-2} + \cdots + x^2 + x + 1) = x^n - 1. \tag{1}$$

For more on Euclid and perfect numbers see *The Elements* Book IX Proposition 36. Moving to the 1600s we come across Marin Mersenne (1588-1648) whom Mersenne numbers are named after, see page 40 of [1]. And ever since Mersenne, mathematicians and other enthusiasts have been searching for Mersenne primes and factoring Mersenne composites. (A composite number is a number which is not prime.) The story of both endeavors is interesting and entertaining. For example, currently the largest known prime is the Mersenne prime $2^{82589933} - 1$ [2] which has $24,862,048$ digits! And back in 2012 the group NFS@Home [3] factored $2^{1061} - 1$ which has 320 digits [4]. $2^{1061} - 1$ is the product of two primes, one with 143 digits and the other with 177 digits.

In this paper we focus on the factorization of Mersenne composites $2^n - 1$ where n is prime. For if n is composite then $\exists\ a, b \in \mathbb{N}$ such that $n = ab$ and therefore

$$x^n - 1 = x^{ab} - 1 = (x^a)^b - 1 = (x^a - 1)(x^{a(b-1)} + x^{a(b-2)} + \cdots + x^{2a} + x^a + 1) \tag{2}$$

by (1). Plugging in $x = 2$ to (2) gives us a nontrivial factor of $2^n - 1$, namely $2^a - 1$. For example no calculations are needed to conclude that $2^{15} - 1$ is composite since $2^3 - 1 = 7$ and $2^5 - 1 = 31$ are both factors. Indeed $2^{15} - 1 = 32767 = 7 \cdot 31 \cdot 151$. In fact there is a more general principle lying around here.

$$x^n - 1 = \prod_{d|n} \Phi_d(x)$$

where the product above is running through all of the divisors, d, of n and $\Phi_d(x)$ is the d^{th} cyclotomic polynomial. The cyclotomic polynomials can be defined recursively as $\Phi_1(x) = x - 1$ and

$$\Phi_m(x) = \frac{x^m - 1}{\prod_{\substack{d|m \\ d<m}} \Phi_d(x)}.$$

Now we can see why $2^{15} - 1$ has three factors since

$$\begin{aligned} x^{15} - 1 &= \Phi_1(x)\Phi_3(x)\Phi_5(x)\Phi_{15}(x) \\ &= (x-1)(x^2 + x + 1)(x^4 + x^3 + x^2 + x + 1)(x^8 - x^7 + x^5 - x^4 + x^3 - x + 1). \end{aligned}$$

It is important to note that for composite n, even though certain factors of $2^n - 1$ are known, it may still be extremely difficult to fully factor $2^n - 1$. To illustrate this fact we can turn to the Cunningham Project [5] which seeks to factor numbers of the form $b^n \pm 1$. On the Cunningham Project's website there is a "wanted" list [6] and the first number on that list is the 364 digit number $2^{1207} - 1$. Since $1207 = 17 \cdot 71$ we know that

$$2^{1207} - 1 = \Phi_1(2)\Phi_{17}(2)\Phi_{71}(2)\Phi_{1207}(2)$$

and

$$\begin{aligned}\Phi_1(2) &= 1\\ \Phi_{17}(2) &= 131071\\ \Phi_{71}(2) &= 228479 \cdot 48544121 \cdot 212885833\end{aligned}$$

however no factor of the 337 digit composite number $\Phi_{1207}(2)$ is known.

Now let p be prime, then $2^p - 1$ may be prime. From above we see that this is a necessary condition, but it is not sufficient. The first example is $2^{11} - 1 = 2047 = 23 \cdot 89$. Note that (1) is of no use in this case since, for a prime p, the polynomial $x^{p-1} + \cdots + x + 1$ is irreducible over $\mathbb{Z}[x]$ by Eisenstein's criterion, see example 4 on page 310 of [7]. However there are some interesting polynomial identities that can be used to factor $2^{11} - 1$. Finding these identities is harder than factoring $2^{11} - 1$, but they are beautiful and lead us to an interesting conjecture. The idea is to find polynomials $f, g \in \mathbb{Z}[x]$ such that $f(2) = 23$ and $g(2) = 89$, then the polynomial $x^{11} - 1 - f(x)g(x)$ will have 2 as a root since $2^{11} - 1 - 23 \cdot 89 = 0$. Hence $\exists\ m \in \mathbb{Z}[x]$ such that

$$x^{11} - 1 - f(x)g(x) = (x - 2)m(x). \tag{3}$$

For different choices of f and g we can see how "nice" or "ugly" the polynomial $m(x)$ can get. First let's start off with some bad examples. Perhaps the simplest polynomials with $f(2) = 23$ and $g(2) = 89$ are $f(x) = x + 21$ and $g(x) = x + 87$ so let's try those. Using Mathematica we find that

$$\begin{aligned}m(x) &= \frac{x^{11} - 1 - (x + 21)(x + 87)}{x - 2}\\ &= x^{10} + 2x^9 + 4x^8 + 8x^7 + 16x^6 + 32x^5 + 64x^4 + 128x^3 + 256x^2 + 511x + 914.\end{aligned}$$

Note that this $m(x)$ contains 11 terms and its coefficients grow exponentially. Now let's try the "binary representation polynomials" of 23 and 89.

$$\begin{aligned}23 = 2^4 + 2^2 + 2 + 1 &\to f(x) = x^4 + x^2 + x + 1\\ 89 = 2^6 + 2^4 + 2^3 + 1 &\to g(x) = x^6 + x^4 + x^3 + 1\\ \Rightarrow m(x) &= x^{10} + x^9 + 2x^8 + 2x^7 + 2x^6 + 2x^5 + 2x^4 + x^3 + x^2 + x + 1\end{aligned}$$

i.e., we have the identity

$$x^{11}-1-(x^4+x^2+x+1)(x^6+x^4+x^3+1) = (x-2)(x^{10}+x^9+2x^8+2x^7+2x^6+2x^5+2x^4+x^3+x^2+x+1).$$

As one more bad example, this one being more random, consider

$$\begin{aligned}f(x) &= 2x^2 + 6x + 3\\ g(x) &= x^6 + 3x^2 + 4x + 5\\ \Rightarrow m(x) &= x^{10} + 2x^9 + 4x^8 + 6x^7 + 6x^6 + 9x^5 + 18x^4 + 30x^3 + 34x^2 + 25x + 8.\end{aligned}$$

Now for some good examples which is the reason why I am writing this. I was surprised to find the following

$$x^{11} - 1 - (x^5 - x^4 + x^3 - x + 1)(x^6 + x^5 - x^3 + 1) = (x - 2)(x^5 + 1). \tag{4}$$

Notice how small $m(x) = x^5 + 1$ is, both in the number of terms and coefficients. I found this identity by trying different m's over $\mathbb{F}_2[x]$ and was happy to discover that this factorization holds over $\mathbb{Z}[x]$. For the purposes of factoring $2^{11} - 1$ it is better to view (3) as

$$x^{11} - 1 - (x - 2)m(x) = f(x)g(x).$$

It turns out that similar identities exist for other prime numbers such as 23, 29, and 37 which motivates the following definition.

Definition: Let n be prime such that $2^n - 1$ is not prime and define the set

$$PM_n = \{m(x) = a_{n-1}x^{n-1} + \cdots + a_1x + a_0 |\ a_i \in \{0,1\} \forall i \text{ and } x^n - 1 - (x-2)m(x) \text{ is reducible}\}.$$

From (4) we see $x^5+1 \in PM_{11}$ and one can perform a brute force search to find that $|PM_{11}| = 54$. Recall that we are checking all 2^{11} polynomials, $m(x)$, with coefficients of either zero or one and $\deg(m) \leq 10$. Below are results for other primes.

n	$\lvert PM_n \rvert$	min $m(1)$	max $m(1)$	mean $m(1)$
11	54	2	9	$\frac{91}{18} \approx 5.056$
23	658	3	17	$\frac{3417}{329} \approx 10.386$
29	1875	4	24	$\frac{24973}{1875} \approx 13.319$
37	–	6	–	–

Table 1: Data on PM_n

Here the min, max, and mean are taken over all $m \in PM_n$, i.e.,

$$\min m(1) = \min_{m \in PM_n} m(1) \qquad \max m(1) = \max_{m \in PM_n} m(1) \qquad \text{mean } m(1) = \frac{1}{|PM_n|} \sum_{m \in PM_n} m(1).$$

Note that $m(1)$ is equal to the number of nonzero terms since each $m(x)$ only has coefficients of zero or one. I am still running the search to find $|PM_{37}|$, however I know that there are at least 10 polynomials in this set. The following is the identity that $x^{34} + x^{31} + x^{22} + x^{19} + x^8 + 1 \in PM_{37}$ produces

$$\begin{aligned} x^{37}&-1 - (x-2)(x^{34} + x^{31} + x^{22} + x^{19} + x^8 + 1) \\ &= (x^8 - x^6 + x^5 - x + 1)(x^{29} + x^{26} + x^{23} + x^{22} - x^{18} + x^{14} + x^{10} + x^8 - x^5 + 1). \end{aligned}$$

My question is can we always find such identities?

Conjecture 1: Let n be prime such that $2^n - 1$ is not prime, then $PM_n \neq \emptyset$, where PM_n is defined above.

And if conjecture 1 is true, does $|PM_n|$ always increase? From table 1 we see that $|PM_{11}| < |PM_{23}| < |PM_{29}|$, does this pattern continue?

Conjecture 2: (Assuming conjecture 1 is true.) Let $n_1 < n_2$ be primes such that $2^{n_1} - 1$ and $2^{n_2} - 1$ are both composite, then $|PM_{n_1}| < |PM_{n_2}|$.

Results

Obviously I cannot list all the identities that I have found here, see my GitHub for that. So instead I will list my favorites according to the following ranking system.

1. the most interesting identities are the ones that fully factor the Mersenne number which is always the case when there are exactly two prime factors but if there are three or more, as in $2^{29} - 1 = 233 \cdot 1103 \cdot 2089$, then the polynomial identity may only give one prime factor.

2. The less terms $m(x)$ has the better. However, if there are more than two polynomials in PM_n achieving $\min m(1)$ then I will just list the one which takes the least characters to write.

3. Factors whose coefficients are in the set $\{-1, 0, 1\}$ are preferred. For example, $x^5+1, x^8+x^3 \in PM_{11}$ yielding

$$x^{11} - 1 - (x-2)(x^5+1) = (x^5 - x^4 + x^3 - x + 1)(x^6 + x^5 - x^3 + 1)$$
$$x^{11} - 1 - (x-2)(x^8+x^3) = (x^4 + x^3 - 1)(x^7 - x^6 + 2x^4 - x^3 + 1)$$

so the first equation is preferred since the polynomials on the right hand side have coefficients in the set $\{-1, 0, 1\}$, while the last polynomial in the second equation has the monomial $2x^4$.

Now according to my rankings here are the most interesting identities that I've found.

Identity for $2^{11} - 1 = 23 \cdot 89$

$$x^{11} - 1 - (x-2)(x^5+1) = (x^5 - x^4 + x^3 - x + 1)(x^6 + x^5 - x^3 + 1)$$

Identities for $2^{23} - 1 = 47 \cdot 178481$

$$\begin{aligned} x^{23} - 1 - (x-2)(x^{15} + x^8 + x^5) \\ = (x^5 + x^4 - 1)(x^{18} - x^{17} + x^{16} - x^{15} + x^{14} - x^{12} + x^{11} + x^9 - x^8 + x^6 - x^5 + x^4 + 1) \\ x^{23} - 1 - (x-2)(x^{15} + x^{10} + x^4) \\ = (x^5 + x^4 - 1)(x^{18} - x^{17} + x^{16} - x^{15} + x^{14} - x^{12} + x^{11} + x^9 - x^8 + 2x^5 - x^4 + 1) \end{aligned}$$

Identities for $2^{29} - 1 = 233 \cdot 1103 \cdot 2089$

Of the 1875 polynomials in PM_{29} only one yields an identity that gives all 3 prime factors of $2^{29} - 1$ making it unique amongst the other 1874.

Let $m(x) = x^{28} + x^{27} + x^{26} + x^{22} + x^{21} + x^{20} + x^{19} + x^{17} + x^{16} + x^{15} + x^{13} + x^{12} + x^8 + x^7 + x^3$, then

$$x^{29} - 1 - (x-2)m(x) = (x^8 - x^5 + x^4 - x^3 + 1)(x^9 + x^8 + 2x^7 + x^6 + x^4 - 1)(x^{11} + x^5 + x^4 - x^3 + 1).$$

The identities with the "shortest" $m(x) \in PM_{29}$ are below.

$$\begin{aligned} x^{29} - 1 - (x-2)(x^{22} + x^{13} + x^{12} + x^8) \\ = (x^{11} + x^6 - x^5 + x^4 - x^3 + x - 1)(x^{18} - x^{13} + x^{11} + x^{10} - x^5 + x^2 + x + 1) \\ x^{29} - 1 - (x-2)(x^{22} + x^{15} + x^{10} + x) \\ = (x^8 - x^5 + x^3 + x - 1)(x^{21} + x^{18} - x^{16} + x^{14} - x^{13} + x^{11} + x^7 + x^6 - x^5 + x^3 - x + 1) \end{aligned}$$

Identity for $2^{37} - 1 = 223 \cdot 616318177$

$$\begin{aligned} x^{37} - 1 - (x-2)(x^{34} + x^{31} + x^{22} + x^{19} + x^8 + 1) \\ = (x^8 - x^6 + x^5 - x + 1)(x^{29} + x^{26} + x^{23} + x^{22} - x^{18} + x^{14} + x^{10} + x^8 - x^5 + 1) \end{aligned}$$

Notice that most of the factors f and g, of $x^n - 1 - (x-2)m(x) = f(x)g(x)$, in the identities above have coefficients in the set $\{-1, 0, 1\}$. Is this always possible?

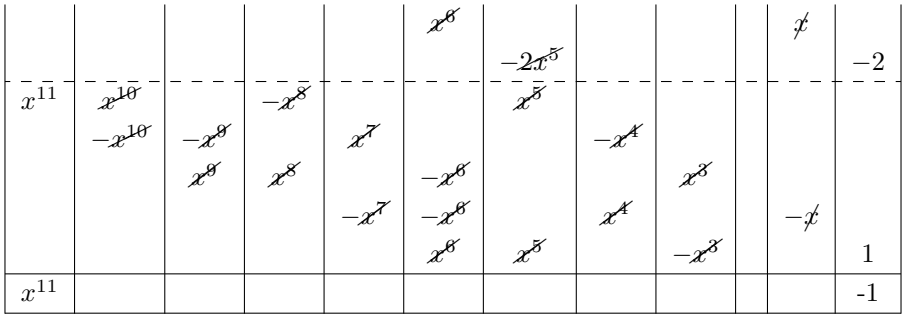
Conjecture 3: Let n be prime such that $2^n - 1$ is not prime, then there exists $m \in PM_n$ and $f, g \in \mathbb{Z}[x]$ with coefficients in $\{-1, 0, 1\}$ such that $x^n - 1 - (x-2)m(x) = f(x)g(x)$.

Final Remarks

As I said in the beginning, finding these polynomial identities is harder than factoring the corresponding the Mersenne number, but perhaps there is some underlying structure that I am unaware of. From number theory we know that if p is prime, then any factor of $2^p - 1$ is of the form $2pk + 1$

for some $k \in \mathbb{N}$. However, I don't know if this fact can be incorporated into these polynomial identities. In any case, the thing that still amazes me is that such small m's can be found. All of the cancellations remind me of the cyclotomic polynomials. As one last example I will leave you with this visual aid of polynomial multiplication.

$$(x-2)(x^5+1)+(x^5-x^4+x^3-x+1)(x^6+x^5-x^3+1)=x^{11}-1$$



References

[1] Stillwell, J. Mathematics and Its History. 3rd edn. Undergraduate Texts in Mathematics. Springer, New York (2010)

[2] Mersenne Primes https://www.mersenne.org/primes/

[3] NFS@Home Website https://escatter11.fullerton.edu/nfs/

[4] Childers, G. Factorization of a 1061-bit number by the Special Number Field Sieve. IACR Cryptol. ePrint Arch. 2012, 444

[5] Cunningham Project Website https://homes.cerias.purdue.edu/~ssw/cun/

[6] Cunningham Project Wanted List https://homes.cerias.purdue.edu/~ssw/cun/want141.txt

[7] Dummit, D.S. and Foote, R.M. (2004) Abstract Algebra. 3rd Edition, John Wiley & Sons, Inc.