As it was reported in the 2004/1 issue of KöMaL, the above sum will be paid as a prize by the Electronic Frontier Foundation (EFF) to the one who first discovers a prime number of ten million digits or more. The largest prime known as yet[1] is $2^{20\,996\,011}-1$. It has 6 320 430 digits. It was found in the Great Internet Mersenne Prime Search (GIMPS) project on 17 November 2003. This programme is open to anyone who wants to join, the details can be found on the web page www.mersenne.org.

Euclid knew them

Mersenne primes are the primes of the form 2^k-1. They are named after *Martin Mersenne* the great French „science organizer" of the 17th century. He was in intense correspondence with Fermat, Descartes and other leading scholars. However, these primes appeared as early as in the search for perfect numbers by the ancient Greeks. Perfect numbers are those that are equal to the sum of their proper factors. Such a number is 6 or 496. Theorem IX.36 of Euclid's book *Elements* states:

If a geometric progression is formed starting at unity with a ratio of two until the sum of the series is a prime, and the sum is multiplied by the last term, the product will be a perfect number.

That is, if $1+2+2^2+\cdots+2^{k-1}=2^k-1$ is a prime then $2^{k-1}(2^k-1)$ is a perfect number. For example, with $k=2$ we get 6, and with $k=5$ we get 496.

To prove (that is roughly what Euclid did, too), the factors of the number $n=2^{k-1}(2^k-1)$ should be added, where $q=2^k-1$ is a prime:

$$1+2+4+\cdots+2^{k-1}+q+2q+\cdots+2^{k-2}q=2^k-1+q(2^{k-1}-1)=q2^{k-1}=n.$$

The perfect number that Euclid's formula gives are all even. It still remains an open problem whether there exists an odd perfect number at all (probably not). On the other hand, Euclid's algorithm generates all even perfect numbers, as shown by Euler 2000 years later. Thus there is a one-to-one correspondence between even perfect numbers and primes of the form 2^k-1. Unfortunately, we still do not know whether the number of such primes is finite or infinite. (Most suspect that the latter is the case.) As Paul Erdős put it, „this question is perhaps the hardest, though not the most pressing problem faced by humankind."

The mysterious list

It was also in connection to perfect numbers that Mersenne investigated primes of the above form. In 1644, he put forward his famous list stating that 2^k-1 is a prime if $k=2,3,5,7,13,17,19,31,67,127$ or 257 but a composite number for all other k less than 257.

Looking at the list, one immediately realizes that there are only prime indices. This is no coincidence, since if k is a composite number, that is $k=uv$, where $u,v>1$, then the identity $a^n-b^n=(a-b)(a^{n-1}+a^{n-2}b+\cdots+b^{n-1})$ can be applied to $2^k-1=(2^u)^v-1^v$. It is divisible by 2^u-1, so 2^k-1 is a composite number, too.

For small values of k it is easy to check whether 2^k-1 is a prime or not, but it becomes harder pretty soon. Even the testing of $2^{31}-1$ takes very long if we do it by the method of trying the integers (larger than 1) up to the square root of the number. (It is enough to try the primes but that requires that we know the primes up to the given limit.) As Mersenne himself wrote, „to decide whether a 15 or 20-digit-number is a prime, a whole lifetime is not enough." That is why it is surprising that he took the courage to compile a list like that (based on considerations that are still not fully known), and it is even more surprising that his list contains only five errors: as it was shown 300(!) years later, $2^{67}-1$ and $2^{257}-1$ are in fact composite numbers while the primes $2^{61}-1$, $2^{89}-1$ and $2^{107}-1$ are missing.

First things first

Of course, Mersenne could have been aware of some results that would have made it easier to find the prime factors of a number of the form $M_p=2^p-1$, where p is a prime. (Such numbers are referred to as Mersenne numbers in the considerations below.) One theorem states that (for $p>2$) each prime factor of M_p is $2kp+1$ and also of the form $8j\pm1$. For example, the prime factors of $M_{43}=2^{43}-1$ are of the forms $86k+1$ and $8j\pm1$, and the smallest such prime, 431 is indeed a factor of M_{43}. Similarly, in order to see that M_{31} is a prime, it is enough to make sure that it is not divisible by any prime of the forms $248t+1$ or $248t+63$. This may be a reason for the index 31 being on the list and 43 not being there (but it still does account for the guessing of most of the numbers on the list.)

Now we will prove that the prime factors of M_p are all of the form $2kp+1$. The proof is based on the concept of *order*.

[1] The article was written in the fall of 2003.

Let c and m be relative primes, and consider the remainders of the powers of c, $1 = c^0, c, c^2, c^3, \ldots, c^n, \ldots$ divided by m. Since the number of remainders is finite, there will be two powers c^i and c^j $(0 \le i < j)$ that have the same remainder, that is, m divides $c^j - c^i = c^i(c^{j-i} - 1)$. Since $(c, m) = 1$, it follows that $c^{j-i} - 1$ is also divisible by m, that is the remainder of c^{j-i} is 1. Let r be the smallest positive integer such that the remainder of c^r is 1. Then the remainders of $1 = c^0, c, c^2, c^3, \ldots, c^n, \ldots$ form a periodic sequence in which the length of the (shortest) period is r. That number r is defined as the order of the number c modulo m and denoted by $o_m(c)$, which is read out as „ordo m c".

We will also make use of Fermat's little theorem. It states that the remainder of c^{q-1} divided by the prime q is 1, provided that q is not a factor of c. It follows from the previous paragraph that the order of c divides $q - 1$, that is $o_q(c) \mid q - 1$.

Now let q be a prime factor of the Mersenne number $M_p = 2^p - 1$, where $p > 2$ is a prime. Then the remainder of 2^p divided by q is 1. Thus $o_q(2) \mid p$. Since p is a prime and the remainder of 2^1 is not 1, $o_q(2)$ can only be p. Hence, according to the previous paragraph, it follows that $p \mid q - 1$ and since $q - 1$ is even that means $q = 2kp + 1$.

The concept of the order and the proof above is easier to formulate in terms of congruences. $a \equiv b \pmod m$ means that a and b give the same remainder when divided by m, that is $m \mid a - b$. The order of $c \bmod m$ is the smallest positive integer r for which $c^r \equiv 1 \pmod m$. Fermat's little theorem states that

$$c^{q-1} \equiv 1 \pmod q$$

if q is a prime and c is not divisible by q, that is $c^q \equiv c \pmod q$ for all c. The property of the order being the length of a period means that

$$c^i \equiv c^j \pmod m \quad \Longleftrightarrow \quad i \equiv j \pmod{o_m(c)}.$$

The statement that the prime factors of M_p are of the form $8j \pm 1$ can be proved by using the theory of quadratic residues, with the help of the so-called Legendre symbols.

The perennial test

It was *Edouard Lucas* in 1876 who made the first breach in the correctness of Mersenne's list. He introduced a completely different method that remains in use to this day. The more than 200 000 computers linked together in a network for the GIMPS project also use his test to search for Mersenne primes. The test is based on the recurrence $a_1 = 4$, $a_{n+1} = a_n^2 - 2$: For a prime $p > 2$, $M_p = 2^p - 1$ is a prime if and only if $M_p \mid a_{p-1}$.

For example, if $p = 7$ then $a_1 = 4$, $a_2 = 14$, $a_3 = 194 \equiv -60 \pmod{127}$, $a_4 \equiv 3598 \equiv 42 \pmod{127}$, $a_5 \equiv 1762 \equiv -16 \pmod{127}$, $a_6 \equiv 254 \equiv 0 \pmod{127}$ therefore, $M_7 = 127$ is a prime.

The above example is only an illustration, it is for large indices that the method is really powerful. Lucas used this test in 1876 to show that M_{67} was a composite number, without being able to present a single factor. It was more than twenty-five years later that Cole managed to factor M_{67}. Lucas also proved that M_{127} was indeed a prime, and that remained the largest prime known until the advent of computers.

As seen in the illustration above, it is not necessary to calculate the numbers a_i themselves, it is enough to consider their remainders when divided by M_p. The remainder is very simple to find with the help of a computer since in binary notation M_p consists of all ones. Thus the task is similar to finding the remainder of a number in decimal notation, say 21 357 246 divided by 999: since the remainder of 10^3 and all 10^{3k} is always 1,

$$21\,357\,246 = 21 \cdot 10^6 + 357 \cdot 10^3 + 246 \equiv 21 + 357 + 246 \equiv 624 \pmod{999},$$

that is the remainder is obtained by simply shifting certain sequences of digits.

A taste of another number system

We will show the sufficiency of the test: *If*

$$M_p \mid a_{p-1}, \tag{1}$$

then M_p is a prime.

(The proof of the condition being necessary uses similar techniques, and it also requires the Legendre symbol mentioned before.)

In the proof, we will use the basic properties of the notions of divisibility, congruency and order introduced for numbers of the form $a + b\sqrt{3}$ (where a, b are integers). These work in the same way as in the set of integers.

It is easy to show by induction that $a_k = \left(2 + \sqrt{3}\right)^{2^{k-1}} + \left(2 - \sqrt{3}\right)^{2^{k-1}}$. Hence condition (1) is equivalent to the divisibility

$$M_p \,\Big|\, \left(2 + \sqrt{3}\right)^{2^{p-2}} + \left(2 - \sqrt{3}\right)^{2^{p-2}}.$$

Multiplying the right-hand side by $\left(2+\sqrt{3}\right)^{2^{p-2}}$ and using $\left(2-\sqrt{3}\right)\left(2+\sqrt{3}\right)=1$, we have

$$M_p \,\Big|\, \left(2+\sqrt{3}\right)^{2^{p-1}}+1, \quad \text{that is} \quad \left(2+\sqrt{3}\right)^{2^{p-1}} \equiv -1 \pmod{M_p}. \tag{2}$$

In order to deduce from (2) that M_p is a prime we need the following **lemma:**

If $q>3$ is an arbitrary prime, then

$$\left(a+b\sqrt{3}\right)^q \equiv a+b\sqrt{3} \quad \text{or} \quad a-b\sqrt{3} \pmod{q}. \tag{3}$$

Proof for the lemma: From the binomial theorem,

$$\left(a+b\sqrt{3}\right)^q = a^q + \binom{q}{1}a^{q-1}b\sqrt{3} + \binom{q}{2}a^{q-2}3b^2 + \cdots + b^q 3^{\frac{q-1}{2}}\sqrt{3}. \tag{4}$$

According to Fermat's little theorem, $a^q \equiv a \pmod{q}$ and $b^q \equiv b \pmod{q}$ and since q is a prime, $\binom{q}{1}, \binom{q}{2}, \ldots, \binom{q}{q-1}$ are all divisible by q. Finally, by virtue of Fermat's little theorem again,

$$q \,\Big|\, \left(3^{\frac{q-1}{2}}\right)^2 - 1 = \left(3^{\frac{q-1}{2}}-1\right)\left(3^{\frac{q-1}{2}}+1\right),$$

and since q is a prime it will always divide one of the factors, that is $3^{\frac{q-1}{2}} \equiv \pm 1 \pmod{q}$. Lemma (3) now follows if these congruences are substituted into (4).

Now we can resume the proof of the theorem: assume that (2) is true and let q be a prime factor of M_p. (It is clear that $q>3$ here.) We need to show that $q=M_p$. Then the congruency mod q in (2) is also true, that is

$$\left(2+\sqrt{3}\right)^{2^{p-1}} \equiv -1 \pmod{q}. \tag{5}$$

By squaring, we have

$$\left(2+\sqrt{3}\right)^{2^{p}} \equiv 1 \pmod{q}. \tag{6}$$

It follows from (5), (6) and the properties of the order that $o_q\left(2+\sqrt{3}\right) \,\Big|\, 2^p$ but $o_q\left(2+\sqrt{3}\right) \nmid 2^{p-1}$, that is $o_q\left(2+\sqrt{3}\right)=2^p$.

On the other hand, it follows from (3) that $\left(2+\sqrt{3}\right)^q \equiv 2 \pm \sqrt{3} \pmod{q}$. If the $+$ sign is valid here, then

$$\left(2+\sqrt{3}\right)^{q-1} = \left(2-\sqrt{3}\right)\left(2+\sqrt{3}\right)^q \equiv \left(2-\sqrt{3}\right)\left(2+\sqrt{3}\right) = 1 \pmod{q},$$

and thus $o_q\left(2+\sqrt{3}\right) = 2^p \le q-1$, which is impossible since $q \le M_p = 2^p-1$.

If the $-$ sign is valid, then it follows similarly that

$$\left(2+\sqrt{3}\right)^{q+1} \equiv 1 \pmod{q},$$

so $o_q\left(2+\sqrt{3}\right) = 2^p \le q+1$. Since $q \le M_p = 2^p-1$, it follows that $q=M_p$, that is M_p is a prime, indeed.

Who is going to win?

There is a good chance that the one hundred thousand dollars of EFF will be awarded for a Mersenne prime, though other competitors are coming up, too. The list of the largest primes known on 17 January 2004 is lead by three Mersenne primes, but the fourth place is taken by $5359 \cdot 2^{5\,054\,502}+1$ (a number of more than one and a half million digits) found in December(!) 2003. Such numbers of the form $r \cdot 2^k+1$, with a little luck, are also relatively easy to test if r is a small odd number. In addition, primes of this type probably occur more frequently than Mersenne primes, so it may happen that a number of at least ten million digits of this kind will be found sooner than a Mersenne prime. However, Mersenne primes establish a wonderful interconnection of more than two thousand years of mathematics, leaving enough work to be done for another two thousand years, and the real winners may be those who are able to contribute to theory as well.