5 Primes

5.1 The Distribution of the Set of Primes

Given the usefulness of primes as the 'building blocks' of the integers, we naturally want to investigate how they are distributed: we'd like answers to questions such as the following.

1. How many primes are there?
2. How many primes are there with a certain property? (e.g. congruent to 3 modulo 4)
3. If we have discovered the first n primes, how much larger is the next?
4. Can we write every even integer ≥ 4 as a sum of two primes?
5. Are there infinitely many primes p such that $p+2$ is also prime?
6. Does there exist at least one prime between any consecutive squares?
7. Are there infinitely many primes of the form N^2+1?

The first three questions can, more or less, be answered, whereas the remaining four are famous conjectures (the Goldbach, Twin Prime, Legendre's and N^2+1 conjectures respectively) that have remained unsolved for over a century.[1]

The first question has the oldest answer: we earlier saw Euclid's Theorem stating that there are infinitely many primes. We can extend his approach to other situations. For example, it is clear that any prime $p \geq 3$ cannot be even and must therefore be congruent to 1 or 3 modulo 4. Consider the following table of the primes p such that $3 \leq p \leq 120$, arranged by remainder modulo 4:

$p \equiv 1 \pmod 4$	5	13	17	29	37	41	53	61	73	89	97	101	109	113	$\cdots$	
$p \equiv 3 \pmod 4$	3	7	11	19	23	31	43	47	59	67	71	79	83	103	107	$\cdots$

It appears that the primes are fairly evenly distributed between the two classes, and we might reasonably conjecture that there are infinitely many primes of each type. This is indeed the case.

Theorem 5.1. *Infinitely many primes are congruent each to 1 and 3 modulo 4.*

Proof of half the Theorem. We modify Euclid's proof. Suppose that there are finitely many primes congruent to 3 modulo 4: list them as $3, p_1, \ldots, p_n$ and define

$$\Pi := 4p_1p_2p_3\cdots p_n + 3$$

Certainly $\Pi \equiv 3 \pmod 4$ and therefore odd, so all primes dividing it are odd. Note that

$$x, y \equiv 1 \pmod 4 \implies xy \equiv 1 \pmod 4 \qquad (*)$$

hence, if all primes dividing Π were congruent to 1, so also would be Π. Plainly Π is divisible by some prime $p \equiv 3 \pmod 4$. By assumption we have all of these, and there are two possibilities:

1. $p = 3$ from which $3 \mid 4p_1p_2p_3\cdots p_n \implies 3 \mid p_i \implies p_i = 3$ for some i; a contradiction.
2. $p = p_i$ for some i, in which case $p \mid 3 \implies p = 3$; again a contradiction.

■

[1] Several results which are very close to these have been proved recently, for example the weak Goldbach conjecture states that every odd integer ≥ 9 is the sum of three odd primes was proved in 2013.

Before moving on, consider why the proof cannot be modified to show that infinitely many primes are congruent to 1 modulo 4. One issue is that the corresponding proposition to $(*)$ is false: in fact

$$x, y \equiv 3 \pmod 4 \implies xy \equiv 1 \pmod 4!$$

and we cannot therefore claim that any $\Pi \equiv 1$ (or $\equiv 3$) is divisible by a prime congruent to 1. Indeed:

- $\Pi := 21 = 3 \cdot 7 \equiv 1 \pmod 4$ is not divisible by any primes congruent to 1.
- $\Pi := 3 \cdot 7 \cdot 11 = 231 \equiv 3 \pmod 4$ is not divisible by any primes congruent to 1.

A simple proof of the $\equiv 1$ part of the Theorem will be given later using quadratic residues.
In fact a much harder and more general result is available.

Theorem 5.2 (Dirichlet). *If* $\gcd(a, m) = 1$, *then infinitely many primes* p *satisfy* $p \equiv a \pmod m$.

Counting Primes Now we turn to the third in our list of questions. To think about this, we introduce the concept of a *counting function*: a function $f : \mathbb{N} \to \mathbb{N}_0$ for which $f(x)$ is the number of positive integers less or equal to x satisfying some property. Euler's totient function φ is an example:

$$\varphi(x) = |\{n \in \mathbb{N}_{\leq x} : \gcd(x, n) = 1\}|$$

Here is another.

Example 5.3. Consider the counting function

$$f(x) = |\{n \in \mathbb{N}_{\leq x} : n \equiv 4 \pmod 7\}|$$

To get a feel for f, compute the first few values:

x	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
$f(x)$	0	0	0	1	1	1	1	1	1	1	2	2	2	2	2	2	2	3	3	3

It seems reasonable to claim that, for large x, $f(x)$ is approximately a seventh of x. More precisely,

$$\frac{x-3}{7} \leq f(x) < \frac{x+4}{7} \implies \frac{x-3}{7x} \leq \frac{f(x)}{x} < \frac{x+4}{7x} \overset{\text{Squeeze}}{\underset{\text{Thm}}{\Longrightarrow}} \lim_{x\to\infty} \frac{f(x)}{x} = \frac{1}{7}$$

There is terminology for this: '$f(x)$ is *asymptotic to* $\frac{1}{7}x$,' and we write

$$f(x) \sim \frac{1}{7}x$$

Intuitively, $f(x)$ *grows like* $\frac{1}{7}x$. This is one way of giving precision to the statement, 'one seventh of the integers are congruent to 4 modulo 7.'

Armed with our new notation, we consider the asymptotic behavior of the primes.

Definition 5.4. $\pi(x) := |\{p : p \leq x\}|$ is the number of primes less than x.

Theorem 5.5 (Prime number theorem). $\pi(x) \sim \frac{x}{\ln x}$. *Otherwise said,* $\lim_{x\to\infty} \frac{\pi(x)}{x/\ln x} = 1$.

A proof is too involved for this course; interpreting the result is tough enough! One approach involves probability: the chance of a random integer in the interval $[1, x]$ being prime is

$$\mathbb{P}\big(y \in [1,x] \text{ prime}\big) = \frac{\pi(x)}{x} \approx \frac{1}{\ln x}$$

While the expression $\frac{x}{\ln x}$ estimates the function $\pi(x)$, it is, in fact, always an under-estimate. A more accurate estimate involves an integral, albeit one that needs its own estimation!

$$\pi(x) \sim \int_2^x \frac{1}{\ln t}\mathrm{d}t$$

Example 5.6. To check the veracity of these claims: consider the 1000^{th} prime $p_{1000} = 7919$:

$$\pi(7919) = 1000, \qquad \frac{7919}{\ln 7919} \approx 882, \qquad \int_2^{7919} \frac{1}{\ln t}\mathrm{d}t \approx 1016$$

A little extra algebra tells us that the n^{th} prime should be located around $p_n \approx n \ln n$. Indeed $1000 \ln 1000 \approx 6908$, which is a 13% under-estimate.

Exercises 5.1 1. (a) Verify that every even number between 70 and 80 is a sum of two primes.

(b) How many different ways can 70 be written as a sum of two primes $70 = p + q$ with $p \leq q$? Repeat the question for 80.

2. (a) Show that if $p \geq 5$ is prime, then $p \equiv \pm 1 \pmod 6$.

(b) Mimic the half-proof of Theorem 5.1 to show that there are infinitely many primes congruent to 5 modulo 6.
(*Hint: let* $\Pi := 6p_1p_2\cdots p_n + 5$ *where* $p_1, \ldots, p_n \equiv 5 \pmod 6$)

3. (a) Explain the statement "one-fifth of all numbers are congruent to 2 modulo 5" by using the counting function

$$F(x) = |\{\text{positive numbers } n \leq x \text{ satisfying } n \equiv 2 \pmod 5\}|$$

(b) Explain the statement "most numbers are not squares" by using the counting function

$$S(x) = |\{\text{square numbers less than } x\}|$$

4. Let n be large. By computing $\frac{x}{\ln x}$ when $x = n \ln n$, argue that $p_n \approx n \ln n$ is a reasonable estimate for the value of the n^{th} prime. Use this expression to argue that, for large n,

$$p_{n+1} - p_n \approx 1 + \ln(n+1)$$

Comment on the values of p_{1000} and p_{1001}.

5. (Hard) Let p be an odd prime and consider the quantity

$$\frac{A_p}{B_p} := 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \cdots + \frac{1}{p-1} \quad \text{where} \quad \gcd(A_p, B_p) = 1$$

(a) Find the value of $A_p \pmod p$ and prove that your answer is correct.

(b) (Even harder - also proves part (a)) Make a conjecture for $A_p \pmod{p^2}$ and prove it.
(*Hint: try adding* $\frac{1}{k} + \frac{1}{p-k}$ *in pairs*)

5.2 Mersenne Primes and Perfect Numbers

While Euclid assures us that the set of primes is infinite, this hasn't prevented a semi-formal competition to find the *largest known prime.* Prior to the advent of computers and mechanical calculators, the largest verified prime had 39 digits. As of early 2022, the largest known prime is $2^{82,589,933} - 1$ with 24,862,048 digits! Such primes have a special name.

Definition 5.7. A *Mersenne prime* is a prime of the form $M_p = 2^p - 1$ where p is itself prime.

These are named for Marin Mersenne, a 17th century French music theorist, mathematician and priest.

Examples 5.8. $M_2 = 2^2 - 1 = 3$, $M_3 = 2^3 - 1 = 7$, $M_5 = 2^5 - 1 = 31$, $M_7 = 2^7 - 1 = 127$. Not all Mersenne numbers are prime, for instance

$$M_{11} = 2^{11} - 1 = 2047 = 23 \cdot 89$$

In fact most Mersenne numbers are not prime; the current largest known prime is only the 51st Mersenne prime to be discovered! It is merely conjectured that there are infinitely many of them.

Whenever the 'world's largest prime' is announced, it is usually a Mersenne prime.[2] There are several reasons for this: a simple motivator is the fact that exponentiation quickly provides large candidates. A related reason is that similar-looking numbers with other bases are never prime:

Theorem 5.9. *If $P = a^n - 1$ is prime for some $a, n \geq 2$, then $a = 2$ and n is prime: that is, P is a Mersenne prime.*

Proof. If $a \geq 3$, then

$$a^n - 1 = (a-1)(a^{n-1} + a^{n-2} + \cdots + a + 1)$$

is composite. By a similar factorization, if $n = mk$ is composite, so also is $2^n - 1$:

$$2^n - 1 = (2^m)^k - 1 = (2^m - 1)((2^m)^{k-1} + (2^m)^{k-2} + \cdots + 1)$$

■

There are many known results about Mersenne primes; look them up if you are interested. We now turn our attention to an old problem which turns out to be related to Mersenne primes, using it partly as an excuse to introduce another commonly-used function.

Definition 5.10. Let $n \in \mathbb{N}$. Define $\sigma(n) = \sum_{d|n} d$ to be the sum of the (positive) divisors of n.

We say that n is *perfect* if it equals the sum of its proper (positive) divisors: that is

$$\sigma(n) = 2n \quad (= \text{proper divisors} + n)$$

Examples 5.11. $6 = 1 + 2 + 3$ and $28 = 1 + 2 + 4 + 7 + 14$ are both perfect.

[2] The Great Internet Mersenne Prime Search is an ongoing collaborative project hunting for such: anyone with a computer can sign up. If you're the first to find a prime with 100 million digits, $100,000 could be yours!

We can compute $\sigma(n)$ similarly to how we evaluated Euler's function. First observe a simple fact following from unique prime factorization:

If $\gcd(m,n)=1$ and $d \mid mn$, then $d = d_1 d_2$ is *uniquely* a product of divisors $d_1 \mid m$ and $d_2 \mid n$

(prime factorization!). When m, n are coprime, it is now immediate that

$$\sigma(mn) = \sum_{d \mid mn} d = \sum_{d_1 \mid m,\, d_2 \mid n} d_1 d_2 = \sum_{d_1 \mid m} d_1 \cdot \sum_{d_2 \mid n} d_2 = \sigma(m)\sigma(n)$$

Moreover, the geometric series formula allows us to easily compute σ applied to a prime power:

$$\sigma(p^\mu) = \sum_{j=0}^{\mu} p^j = \frac{p^{\mu+1}-1}{p-1}$$

and we've now proved the main result:

Theorem 5.12. *σ is multiplicative. Moreover, if $n = p_1^{\mu_1} \cdots p_k^{\mu_k}$ is the prime decomposition of n, then*

$$\sigma(n) = \prod_{j=1}^{k} \frac{p_j^{\mu_j+1}-1}{p_j-1}$$

Examples 5.13. The sum of the positive divisors of $260 = 2^2 \cdot 5 \cdot 13$ is

$$\sigma(260) = \frac{2^3-1}{1} \cdot \frac{5^2-1}{4} \cdot \frac{13^2-1}{12} = 588$$

This can tediously be checked since 260 has divisors 1, 2, 4, 5, 10, 13, 20, 26, 52, 65, 130, 260.
Repeating with $n = 1000 = 2^3 \cdot 5^3$, we see that

$$\sigma(1000) = \frac{2^4-1}{2-1} \cdot \frac{5^4-1}{5-1} = 2340$$

There is an intimate relation between perfect numbers and Mersenne primes: half of it indeed appears in Euclid's *Elements*.

Theorem 5.14. *If $2^p - 1$ is a Mersenne prime, then $2^{p-1}(2^p-1)$ is perfect.*

Proof. Suppose that $M_p = 2^p - 1$ is a Mersenne prime. Since $2^p - 1$ is prime,

$$\sigma(M_p) = \sigma(2^{p-1})\sigma(2^p-1) = \frac{2^p-1}{2-1} \cdot (2^p - 1 + 1) = 2 \cdot 2^{p-1}(2^p-1) = 2M_p$$

■

For small values of p we have the following table: the numbers increase very quickly!

p	$2^p - 1$	$n = 2^{p-1}(2^p - 1)$
2	3	$6 = 1 + 2 + 3$
3	7	$28 = 1 + 2 + 4 + 7 + 14$
5	31	$496 = 1 + 2 + 4 + 8 + 16 + 31 + 62 + 124 + 248$
7	127	$8128 = 1 + 2 + 4 + 8 + 16 + 32 + 64 + 127 + 254 + 508 + 1016 + 2032 + 4064$
13	8191	33,550,336
17	131,072	8,589,869,056

It was conjectured in the middle ages and proved in the 1700's that all *even* perfect numbers have this form.

Theorem 5.15 (Euler). *Every even perfect number has the form $2^{p-1}(2^p - 1)$ for some Mersenne prime $M_p = 2^p - 1$.*

Proof. Suppose that $n = 2^k m$ is an even perfect number, where $k \geq 1$ and m is odd. Our goal is to prove that *m is prime;* we will do this by showing that $\sigma(m) = m + 1$.

Since n is perfect and $\gcd(2^k, m) = 1$, we have two expressions for $\sigma(n)$:

$$\sigma(n) = \begin{cases} 2n = 2^{k+1}m \\ \sigma(2^k)\sigma(m) = (2^{k+1} - 1)\sigma(m) \end{cases} \implies (2^{k+1} - 1)\sigma(m) = 2^{k+1}m$$

Since $2^{k+1} - 1$ is odd, we see that $2^{k+1} \mid \sigma(m)$ so that $\sigma(m) = 2^{k+1}\alpha$ for some $\alpha \in \mathbb{N}$. We now have

$$(2^{k+1} - 1)\alpha = m$$

If we can show that $\alpha = 1$ then we are finished: in such a case

$$\sigma(m) = 2^{k+1} = (2^{k+1} - 1) + 1 = m + 1$$

whence m is prime.

To obtain a contradiction, assume that $\alpha > 1$. Then m is divisible by the distinct divisors $1, \alpha, m$. But then

$$2^{k+1}\alpha = \sigma(m) \geq 1 + \alpha + m = 1 + \alpha + (2^{k+1} - 1)\alpha = 1 + 2^{k+1}\alpha$$

Contradiction!

We conclude that $m = 2^{k+1} - 1$ is prime. By Theorem 5.9 we see that $k + 1 = p$ must also be prime, whence $m = M_p$ is a *Mersenne prime.* ∎

Since only fifty-one Mersenne primes have thus far been discovered, only fifty-one perfect numbers are known to exist, with the currently known largest having 49,724,095 digits! Of course the conjectured infinity of Mersenne primes would also imply the existence of infinitely many even perfect numbers. It remains unknown whether there are any *odd* perfect numbers.

Exercises 5.2 1. Prove that p is prime if and only if $\sigma(p) = p + 1$.

2. Suppose that $M_p = 2^p - 1$ is a Mersenne prime. List all the divisors of $2^{p-1}(2^p - 1)$ and use the geometric sequence formula to explicitly sum them. Hence provide a more explicit proof of Theorem 5.14.

3. Define $\tau(n)$ to be the number of positive divisors of n. Prove that τ is multiplicative and find a formula for $\tau(n)$ in terms of the prime decomposition of $n = p_1^{\mu_1} \cdots p_k^{\mu_k}$. Hence or otherwise, find the number of positive divisors of 1,000,000.

4. If $a^n + 1$ is prime for some integers $a \geq 2$ and $n \geq 1$, show that n must be a power of 2.

   (*Hints: if n is odd, show that $(a+1) \mid (a^n + 1)$ similarly to the proof of Theorem 5.14. Then write $n = 2m$, $a^2 = b$ and repeat...*)

5. Primes of the form $F_k = 2^{2^k} + 1$ are called *Fermat primes*.[3] For instance

   $$F_1 = 5, \quad F_2 = 17, \quad F_3 = 257, \quad F_4 = 65537$$

   (a) If $k \geq 2$, prove that the final digit of F_k is 7.
   (*Hint: Think modulo 2 and 5. What is the period of 2^m modulo 5?*)

   (b) Show that if $k \neq m$, then F_k and F_m are coprime.
   (*Hint: if $k > m$, show that F_m divides $F_k - 2$*)

6. Suppose $n \mid M_p$ where p is an odd prime. Prove that $n = 2kp + 1$ for some integer k.
   (*Hint: if q is a prime divisor of $2^p - 1$, think about why p should divide $q - 1$*)

   The remaining questions consider the potential impossibility of odd perfect numbers.

7. (a) Show that a power of 3 can never be a perfect number.
   (b) More generally, if p is an odd prime, show that p^k is not perfect.

8. (a) Show that a number of the form $3^i 5^j$ can never be perfect.
   (b) More generally, if $p \geq 5$ is an odd prime, show that the product $3^i p^j$ can never be perfect.
   (c) Even more generally, show that if p and q are distinct odd primes, then a number of the form $q^i p^j$ can never be perfect.

9. (Hard) Show that $3^i 5^j 7^k$ is never perfect.
   (*Hint: consider $\sigma(n)$ and sums $1 + 3 + 3^2 + \cdots$, etc. modulo 4, then think modulo 5*)

[3] Fermat thought that all the F_k might be prime, however Euler (1732) and Clausen/Landry (1855/1880) successively showed that F_5 and F_6 are composite with prime factorizations:

$$F_5 = 4{,}294{,}967{,}297 = 641 \cdot 6700417, \qquad F_6 = 18{,}446{,}744{,}073{,}709{,}551{,}617 = 274{,}177 \cdot 67{,}280{,}421{,}310{,}721$$

These were incredible achievements for the time. As of 2022, no other Fermat primes have been discovered, and only up to F_{11} has been completely factored! A distributed computing project similar to GIMPS continues the search...