CHAPTER 3

PRIMES AND THEIR DISTRIBUTION

Mighty are numbers, joined with art resistless.
EURIPIDES

3.1 THE FUNDAMENTAL THEOREM OF ARITHMETIC

Essential to everything discussed herein—in fact, essential to every aspect of number theory—is the notion of a prime number. We have previously observed that any integer $a > 1$ is divisible by ± 1 and $\pm\, a$; if these exhaust the divisors of a, then it is said to be a prime number. In Definition 3.1 we state this somewhat differently.

Definition 3.1. An integer $p > 1$ is called a *prime number*, or simply a *prime*, if its only positive divisors are 1 and p. An integer greater than 1 that is not a prime is termed *composite*.

Among the first 10 positive integers, 2, 3, 5, 7 are primes and 4, 6, 8, 9, 10 are composite numbers. Note that the integer 2 is the only even prime, and according to our definition the integer 1 plays a special role, being neither prime nor composite.

In the rest of this book, the letters p and q will be reserved, so far as is possible, for primes.

Proposition 14 of Book IX of Euclid's *Elements* embodies the result that later became known as the Fundamental Theorem of Arithmetic, namely, that every integer greater than 1 can, except for the order of the factors, be represented as a product of primes in one and only one way. To quote the proposition itself: "If a number be the least that is measured by prime numbers, it will not be measured by any other prime except those originally measuring it." Because every number $a > 1$ is either a prime or, by the Fundamental Theorem, can be broken down into unique prime factors and no further, the primes serve as the building blocks from which all other integers can be made. Accordingly, the prime numbers have intrigued mathematicians through the ages, and although a number of remarkable theorems relating to their distribution in the sequence of positive integers have been proved, even more remarkable is what remains unproved. The open questions can be counted among the outstanding unsolved problems in all of mathematics.

To begin on a simpler note, we observe that the prime 3 divides the integer 36, where 36 may be written as any one of the products

$$6 \cdot 6 = 9 \cdot 4 = 12 \cdot 3 = 18 \cdot 2$$

In each instance, 3 divides at least one of the factors involved in the product. This is typical of the general situation, the precise result being Theorem 3.1.

Theorem 3.1. If p is a prime and $p \mid ab$, then $p \mid a$ or $p \mid b$.

Proof. If $p \mid a$, then we need go no further, so let us assume that $p \nmid a$. Because the only positive divisors of p are 1 and p itself, this implies that $\gcd(p, a) = 1$. (In general, $\gcd(p, a) = p$ or $\gcd(p, a) = 1$ according as $p \mid a$ or $p \nmid a$.) Hence, citing Euclid's lemma, we get $p \mid b$.

This theorem easily extends to products of more than two terms.

Corollary 1. If p is a prime and $p \mid a_1 a_2 \cdots a_n$, then $p \mid a_k$ for some k, where $1 \leq k \leq n$.

Proof. We proceed by induction on n, the number of factors. When $n = 1$, the stated conclusion obviously holds; whereas when $n = 2$, the result is the content of Theorem 3.1. Suppose, as the induction hypothesis, that $n > 2$ and that whenever p divides a product of less than n factors, it divides at least one of the factors. Now $p \mid a_1 a_2 \cdots a_n$. From Theorem 3.1, either $p \mid a_n$ or $p \mid a_1 a_2 \cdots a_{n-1}$. If $p \mid a_n$, then we are through. As regards the case where $p \mid a_1 a_2 \cdots a_{n-1}$, the induction hypothesis ensures that $p \mid a_k$ for some choice of k, with $1 \leq k \leq n - 1$. In any event, p divides one of the integers $a_1, a_2, \ldots, a_n$.

Corollary 2. If $p, q_1, q_2, \ldots, q_n$ are all primes and $p \mid q_1 q_2 \cdots q_n$, then $p = q_k$ for some k, where $1 \leq k \leq n$.

Proof. By virtue of Corollary 1, we know that $p \mid q_k$ for some k, with $1 \leq k \leq n$. Being a prime, q_k is not divisible by any positive integer other than 1 or q_k itself. Because $p > 1$, we are forced to conclude that $p = q_k$.

With this preparation out of the way, we arrive at one of the cornerstones of our development, the Fundamental Theorem of Arithmetic. As indicated earlier, this theorem asserts that every integer greater than 1 can be factored into primes in essentially one way; the linguistic ambiguity *essentially* means that $2 \cdot 3 \cdot 2$ is not considered as being a different factorization of 12 from $2 \cdot 2 \cdot 3$. We state this precisely in Theorem 3.2.

Theorem 3.2 Fundamental Theorem of Arithmetic. Every positive integer $n > 1$ is either a prime or a product of primes; this representation is unique, apart from the order in which the factors occur.

Proof. Either n is a prime or it is composite; in the former case, there is nothing more to prove. If n is composite, then there exists an integer d satisfying $d \mid n$ and $1 < d < n$. Among all such integers d, choose p_1 to be the smallest (this is possible by the Well-Ordering Principle). Then p_1 must be a prime number. Otherwise it too would have a divisor q with $1 < q < p_1$; but then $q \mid p_1$ and $p_1 \mid n$ imply that $q \mid n$, which contradicts the choice of p_1 as the smallest positive divisor, not equal to 1, of n.

We therefore may write $n = p_1 n_1$, where p_1 is prime and $1 < n_1 < n$. If n_1 happens to be a prime, then we have our representation. In the contrary case, the argument is repeated to produce a second prime number p_2 such that $n_1 = p_2 n_2$; that is,

$$n = p_1 p_2 n_2 \qquad 1 < n_2 < n_1$$

If n_2 is a prime, then it is not necessary to go further. Otherwise, write $n_2 = p_3 n_3$, with p_3 a prime:

$$n = p_1 p_2 p_3 n_3 \qquad 1 < n_3 < n_2$$

The decreasing sequence

$$n > n_1 > n_2 > \cdots > 1$$

cannot continue indefinitely, so that after a finite number of steps n_{k-1} is a prime, call it, p_k. This leads to the prime factorization

$$n = p_1 p_2 \cdots p_k$$

To establish the second part of the proof—the uniqueness of the prime factorization—let us suppose that the integer n can be represented as a product of primes in two ways; say,

$$n = p_1 p_2 \cdots p_r = q_1 q_2 \cdots q_s \qquad r \leq s$$

where the p_i and q_j are all primes, written in increasing magnitude so that

$$p_1 \leq p_2 \leq \cdots \leq p_r \qquad q_1 \leq q_2 \leq \cdots \leq q_s$$

Because $p_1 \mid q_1 q_2 \cdots q_s$, Corollary 2 of Theorem 3.1 tells us that $p_1 = q_k$ for some k; but then $p_1 \geq q_1$. Similar reasoning gives $q_1 \geq p_1$, whence $p_1 = q_1$. We may cancel this common factor and obtain

$$p_2 p_3 \cdots p_r = q_2 q_3 \cdots q_s$$

Now repeat the process to get $p_2 = q_2$ and, in turn,

$$p_3 p_4 \cdots p_r = q_3 q_4 \cdots q_s$$

Continue in this fashion. If the inequality $r < s$ were to hold, we would eventually arrive at

$$1 = q_{r+1} q_{r+2} \cdots q_s$$

which is absurd, because each $q_j > 1$. Hence, $r = s$ and

$$p_1 = q_1 \qquad p_2 = q_2, \ldots, p_r = q_r$$

making the two factorizations of n identical. The proof is now complete.

Of course, several of the primes that appear in the factorization of a given positive integer may be repeated, as is the case with $360 = 2 \cdot 2 \cdot 2 \cdot 3 \cdot 3 \cdot 5$. By collecting like primes and replacing them by a single factor, we can rephrase Theorem 3.2 as a corollary.

Corollary. Any positive integer $n > 1$ can be written uniquely in a *canonical form*

$$n = p_1^{k_1} p_2^{k_2} \cdots p_r^{k_r}$$

where, for $i = 1, 2, \ldots, r$, each k_i is a positive integer and each p_i is a prime, with $p_1 < p_2 < \cdots < p_r$.

To illustrate, the canonical form of the integer 360 is $360 = 2^3 \cdot 3^2 \cdot 5$. As further examples we cite

$$4725 = 3^3 \cdot 5^2 \cdot 7 \qquad \text{and} \qquad 17460 = 2^3 \cdot 3^2 \cdot 5 \cdot 7^2$$

Prime factorizations provide another means of calculating greatest common divisors. For suppose that $p_1, p_2, \ldots, p_n$ are the distinct primes that divide either of a or b. Allowing zero exponents, we can write

$$a = p_1^{k_1} p_2^{k_2} \cdots p_n^{k_n}, \qquad b = p_1^{j_1} p_2^{j_2} \cdots p_n^{j_n}$$

Then

$$\gcd(a, b) = p_1^{r_1} p_2^{r_2} \cdots p_n^{r_n}$$

where $r_i = \min(k_i, j_i)$, the smaller of the two exponents associated with p_i in the two representations. In the case $a = 4725$ and $b = 17460$, we would have

$$4725 = 2^0 \cdot 3^3 \cdot 5^2 \cdot 7, \quad 7460 = 2^3 \cdot 3^2 \cdot 5 \cdot 7^2$$

and so

$$\gcd(4725, 17460) = 2^0 \cdot 3^2 \cdot 5 \cdot 7 \cdot = 315$$

This is an opportune moment to insert a famous result of Pythagoras. Mathematics as a science began with Pythagoras (569–500 B.C.), and much of the content of Euclid's *Elements* is due to Pythagoras and his school. The Pythagoreans deserve the credit for being the first to classify numbers into odd and even, prime and composite.

Theorem 3.3 Pythagoras. The number $\sqrt{2}$ is irrational.

Proof. Suppose, to the contrary, that $\sqrt{2}$ is a rational number, say, $\sqrt{2} = a/b$, where a and b are both integers with $\gcd(a, b) = 1$. Squaring, we get $a^2 = 2b^2$, so that $b \mid a^2$. If $b > 1$, then the Fundamental Theorem of Arithmetic guarantees the existence of a prime p such that $p \mid b$. It follows that $p \mid a^2$ and, by Theorem 3.1, that $p \mid a$; hence, $\gcd(a, b) \geq p$. We therefore arrive at a contradiction, unless $b = 1$. But if this happens, then $a^2 = 2$, which is impossible (we assume that the reader is willing to grant that no integer can be multiplied by itself to give 2). Our supposition that $\sqrt{2}$ is a rational number is untenable, and so $\sqrt{2}$ must be irrational.

There is an interesting variation on the proof of Theorem 3.3. If $\sqrt{2} = a/b$ with $\gcd(a, b) = 1$, there must exist integers r and s satisfying $ar + bs = 1$. As a result,

$$\sqrt{2} = \sqrt{2}(ar + bs) = (\sqrt{2}a)r + (\sqrt{2}b)s = 2br + as$$

This representation of $\sqrt{2}$ leads us to conclude that $\sqrt{2}$ is an integer, an obvious impossibility.

PROBLEMS 3.1

1. It has been conjectured that there are infinitely many primes of the form $n^2 - 2$. Exhibit five such primes.

2. Give an example to show that the following conjecture is not true: Every positive integer can be written in the form $p + a^2$, where p is either a prime or 1, and $a \geq 0$.

3. Prove each of the assertions below:

(a) Any prime of the form $3n + 1$ is also of the form $6m + 1$.

(b) Each integer of the form $3n + 2$ has a prime factor of this form.

(c) The only prime of the form $n^3 - 1$ is 7.
[*Hint:* Write $n^3 - 1$ as $(n - 1)(n^2 + n + 1)$.]

(d) The only prime p for which $3p + 1$ is a perfect square is $p = 5$.

(e) The only prime of the form $n^2 - 4$ is 5.

4. If $p \geq 5$ is a prime number, show that $p^2 + 2$ is composite.
[*Hint:* p takes one of the forms $6k + 1$ or $6k + 5$.]

5. (a) Given that p is a prime and $p \mid a^n$, prove that $p^n \mid a^n$.

(b) If $\gcd(a, b) = p$, a prime, what are the possible values of $\gcd(a^2, b^2)$, $\gcd(a^2, b)$ and $\gcd(a^3, b^2)$?

6. Establish each of the following statements:

(a) Every integer of the form $n^4 + 4$, with $n > 1$, is composite.
[*Hint:* Write $n^4 + 4$ as a product of two quadratic factors.]

(b) If $n > 4$ is composite, then n divides $(n - 1)!$.

(c) Any integer of the form $8^n + 1$, where $n \geq 1$, is composite.
[*Hint:* $2^n + 1 \mid 2^{3n} + 1$.]

(d) Each integer $n > 11$ can be written as the sum of two composite numbers.
[*Hint:* If n is even, say $n = 2k$, then $n - 6 = 2(k - 3)$; for n odd, consider the integer $n - 9$.]

7. Find all prime numbers that divide 50!.

8. If $p \geq q \geq 5$ and p and q are both primes, prove that $24 \mid p^2 - q^2$.

9. (a) An unanswered question is whether there are infinitely many primes that are 1 more than a power of 2, such as $5 = 2^2 + 1$. Find two more of these primes.

(b) A more general conjecture is that there exist infinitely many primes of the form $n^2 + 1$; for example, $257 = 16^2 + 1$. Exhibit five more primes of this type.

10. If $p \neq 5$ is an odd prime, prove that either $p^2 - 1$ or $p^2 + 1$ is divisible by 10.

11. Another unproven conjecture is that there are an infinitude of primes that are 1 less than a power of 2, such as $3 = 2^2 - 1$.

(a) Find four more of these primes.

(b) If $p = 2^k - 1$ is prime, show that k is an odd integer, except when $k = 2$.
[*Hint:* $3 \mid 4^n - 1$ for all $n \geq 1$.]

12. Find the prime factorization of the integers 1234, 10140, and 36000.

13. If $n > 1$ is an integer not of the form $6k + 3$, prove that $n^2 + 2^n$ is composite.
[*Hint:* Show that either 2 or 3 divides $n^2 + 2^n$.]

14. It has been conjectured that every even integer can be written as the difference of two consecutive primes in infinitely many ways. For example,

$$6 = 29 - 23 = 137 - 131 = 599 - 593 = 1019 - 1013 = \cdots$$

Express the integer 10 as the difference of two consecutive primes in 15 ways.

15. Prove that a positive integer $a > 1$ is a square if and only if in the canonical form of a all the exponents of the primes are even integers.

16. An integer is said to be *square-free* if it is not divisible by the square of any integer greater than 1. Prove the following:

(a) An integer $n > 1$ is square-free if and only if n can be factored into a product of distinct primes.

(b) Every integer $n > 1$ is the product of a square-free integer and a perfect square. [*Hint:* If $n = p_1^{k_1} p_2^{k_2} \cdots p_s^{k_s}$ is the canonical factorization of n, then write $k_i = 2q_i + r_i$ where $r_i = 0$ or 1 according as k_i is even or odd.]

17. Verify that any integer n can be expressed as $n = 2^k m$, where $k \geq 0$ and m is an odd integer.

18. Numerical evidence makes it plausible that there are infinitely many primes p such that $p + 50$ is also prime. List 15 of these primes.

19. A positive integer n is called *square-full*, or *powerful*, if $p^2 \mid n$ for every prime factor p of n (there are 992 square-full numbers less than 250,000). If n is square-full, show that it can be written in the form $n = a^2 b^3$, with a and b positive integers.

3.2 THE SIEVE OF ERATOSTHENES

Given a particular integer, how can we determine whether it is prime or composite and, in the latter case, how can we actually find a nontrivial divisor? The most obvious approach consists of successively dividing the integer in question by each of the numbers preceding it; if none of them (except 1) serves as a divisor, then the integer must be prime. Although this method is very simple to describe, it cannot be regarded as useful in practice. For even if one is undaunted by large calculations, the amount of time and work involved may be prohibitive.

There is a property of composite numbers that allows us to reduce materially the necessary computations—but still the process remains cumbersome. If an integer $a > 1$ is composite, then it may be written as $a = bc$, where $1 < b < a$ and $1 < c < a$. Assuming that $b \leq c$, we get $b^2 \leq bc = a$, and so $b \leq \sqrt{a}$. Because $b > 1$, Theorem 3.2 ensures that b has at least one prime factor p. Then $p \leq b \leq \sqrt{a}$; furthermore, because $p \mid b$ and $b \mid a$, it follows that $p \mid a$. The point is simply this: a composite number a will always possess a prime divisor p satisfying $p \leq \sqrt{a}$.

In testing the primality of a specific integer $a > 1$, it therefore suffices to divide a by those primes not exceeding $\sqrt{a}$ (presuming, of course, the availability of a list of primes up to $\sqrt{a}$). This may be clarified by considering the integer $a = 509$. Inasmuch as $22 < \sqrt{509} < 23$, we need only try out the primes that are not larger than 22 as possible divisors, namely, the primes 2, 3, 5, 7, 11, 13, 17, 19. Dividing 509 by each of these, in turn, we find that none serves as a divisor of 509. The conclusion is that 509 must be a prime number.

Example 3.1. The foregoing technique provides a practical means for determining the canonical form of an integer, say $a = 2093$. Because $45 < \sqrt{2093} < 46$, it is enough to examine the primes 2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43. By trial, the first of these to divide 2093 is 7, and $2093 = 7 \cdot 299$. As regards the integer 299, the seven primes that are less than 18 (note that $17 < \sqrt{299} < 18$) are 2, 3, 5, 7, 11, 13, 17. The first prime divisor of 299 is 13 and, carrying out the required division, we obtain $299 = 13 \cdot 23$. But 23 is itself a prime, whence 2093 has exactly three prime factors, 7, 13, and 23:

$$2093 = 7 \cdot 13 \cdot 23$$

Another Greek mathematician whose work in number theory remains significant is Eratosthenes of Cyrene (276–194 B.C.). Although posterity remembers him mainly as the director of the world-famous library at Alexandria, Eratosthenes was gifted in all branches of learning, if not of first rank in any; in his own day, he was nicknamed "Beta" because, it was said, he stood at least second in every field. Perhaps the most impressive feat of Eratosthenes was the accurate measurement of the earth's circumference by a simple application of Euclidean geometry.

We have seen that if an integer $a > 1$ is not divisible by any prime $p \leq \sqrt{a}$, then a is of necessity a prime. Eratosthenes used this fact as the basis of a clever technique, called the *Sieve of Eratosthenes*, for finding all primes below a given integer n. The scheme calls for writing down the integers from 2 to n in their natural order and then systematically eliminating all the composite numbers by striking out all multiples $2p, 3p, 4p, 5p, \ldots$ of the primes $p \leq \sqrt{n}$. The integers that are left on the list—those that do not fall through the "sieve"—are primes.

To see an example of how this works, suppose that we wish to find all primes not exceeding 100. Consider the sequence of consecutive integers $2, 3, 4, \ldots, 100$. Recognizing that 2 is a prime, we begin by crossing out all even integers from our listing, except 2 itself. The first of the remaining integers is 3, which must be a prime. We keep 3, but strike out all higher multiples of 3, so that $9, 15, 21, \ldots$ are now removed (the even multiples of 3 having been removed in the previous step). The smallest integer after 3 that has not yet been deleted is 5. It is not divisible by either 2 or 3—otherwise it would have been crossed out—hence, it is also a prime. All proper multiples of 5 being composite numbers, we next remove $10, 15, 20, \ldots$ (some of these are, of course, already missing), while retaining 5 itself. The first surviving integer 7 is a prime, for it is not divisible by 2, 3, or 5, the only primes that precede it. After eliminating the proper multiples of 7, the largest prime less than $\sqrt{100} = 10$, all composite integers in the sequence $2, 3, 4, \ldots, 100$ have fallen through the sieve. The positive integers that remain, to wit, 2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47, 53, 59, 61, 67, 71, 73, 79, 83, 89, 97, are all of the primes less than 100.

The following table represents the result of the completed sieve. The multiples of 2 are crossed out by \; the multiples of 3 are crossed out by /; the multiples of 5 are crossed out by —; the multiples of 7 are crossed out by $\sim$.

	2	3	~~4~~	5	~~6~~	7	~~8~~	~~9~~	~~10~~
11	~~12~~	13	~~14~~	~~15~~	~~16~~	17	~~18~~	19	~~20~~
~~21~~	~~22~~	23	~~24~~	~~25~~	~~26~~	~~27~~	~~28~~	29	~~30~~
31	~~32~~	~~33~~	~~34~~	~~35~~	~~36~~	37	~~38~~	~~39~~	~~40~~
41	~~42~~	43	~~44~~	~~45~~	~~46~~	47	~~48~~	~~49~~	~~50~~
~~51~~	~~52~~	53	~~54~~	~~55~~	~~56~~	~~57~~	~~58~~	59	~~60~~
61	~~62~~	~~63~~	~~64~~	~~65~~	~~66~~	67	~~68~~	~~69~~	~~70~~
71	~~72~~	73	~~74~~	~~75~~	~~76~~	~~77~~	~~78~~	79	~~80~~
~~81~~	~~82~~	83	~~84~~	~~85~~	~~86~~	~~87~~	~~88~~	89	~~90~~
~~91~~	~~92~~	~~93~~	~~94~~	~~95~~	~~96~~	97	~~98~~	~~99~~	~~100~~

By this point, an obvious question must have occurred to the reader. Is there a largest prime number, or do the primes go on forever? The answer is to be found in a remarkably simple proof given by Euclid in Book IX of his *Elements*. Euclid's argument is universally regarded as a model of mathematical elegance. Loosely speaking, it goes like this: Given any finite list of prime numbers, one can always find a prime not on the list; hence, the number of primes is infinite. The actual details appear below.

Theorem 3.4 Euclid. There is an infinite number of primes.

Proof. Euclid's proof is by contradiction. Let $p_1 = 2$, $p_2 = 3$, $p_3 = 5$, $p_4 = 7, \ldots$ be the primes in ascending order, and suppose that there is a last prime, called p_n. Now consider the positive integer

$$P = p_1 p_2 \cdots p_n + 1$$

Because $P > 1$, we may put Theorem 3.2 to work once again and conclude that P is divisible by some prime p. But $p_1, p_2, \ldots, p_n$ are the only prime numbers, so that p must be equal to one of $p_1, p_2, \ldots, p_n$. Combining the divisibility relation $p \mid p_1 p_2 \cdots p_n$ with $p \mid P$, we arrive at $p \mid P - p_1 p_2 \cdots p_n$ or, equivalently, $p \mid 1$. The only positive divisor of the integer 1 is 1 itself and, because $p > 1$, a contradiction arises. Thus, no finite list of primes is complete, whence the number of primes is infinite.

For a prime p, define $p^\#$ to be the product of all primes that are less than or equal to p. Numbers of the form $p^\# + 1$ might be termed *Euclidean numbers*, because they appear in Euclid's scheme for proving the infinitude of primes. It is interesting to note that in forming these integers, the first five, namely,

$$\begin{aligned}
2^\# + 1 &= 2 + 1 = 3 \\
3^\# + 1 &= 2 \cdot 3 + 1 = 7 \\
5^\# + 1 &= 2 \cdot 3 \cdot 5 + 1 = 31 \\
7^\# + 1 &= 2 \cdot 3 \cdot 5 \cdot 7 + 1 = 211 \\
11^\# + 1 &= 2 \cdot 3 \cdot 5 \cdot 7 \cdot 11 + 1 = 2311
\end{aligned}$$

are all prime numbers. However,

$$\begin{aligned}13^{\#}+1&=59\cdot 509\\17^{\#}+1&=19\cdot 97\cdot 277\\19^{\#}+1&=347\cdot 27953\end{aligned}$$

are not prime. A question whose answer is not known is whether there are infinitely many primes p for which $p^{\#}+1$ is also prime. For that matter, are there infinitely many composite $p^{\#}+1$?

At present, 22 primes of the form $p^{\#}+1$ have been identified. The first few correspond to the values $p = 2, 3, 5, 7, 11, 31, 379, 1019, 1021, 2657, 3229$. The twenty-second occurs when $p = 392113$ and consists of 169966 digits. It was found in 2001.

Euclid's theorem is too important for us to be content with a single proof. Here is a variation in the reasoning: Form the infinite sequence of positive integers

$$\begin{aligned}n_1&=2\\n_2&=n_1+1\\n_3&=n_1n_2+1\\n_4&=n_1n_2n_3+1\\&\vdots\\n_k&=n_1n_2\cdots n_{k-1}+1\\&\vdots\end{aligned}$$

Because each $n_k > 1$, each of these integers is divisible by a prime. But no two n_k can have the same prime divisor. To see this, let $d = \gcd(n_i, n_k)$ and suppose that $i < k$. Then d divides n_i and, hence, must divide $n_1n_2\cdots n_{k-1}$. Because $d \mid n_k$, Theorem 2.2 (g) tells us that $d \mid n_k - n_1n_2\cdots n_{k-1}$ or $d \mid 1$. The implication is that $d = 1$, and so the integers $n_k (k = 1, 2, \ldots)$ are pairwise relatively prime. The point we wish to make is that there are as many distinct primes as there are integers n_k, namely, infinitely many of them.

Let p_n denote the nth of the prime numbers in their natural order. Euclid's proof shows that the expression $p_1p_2\cdots p_n+1$ is divisible by at least one prime. If there are several such prime divisors, then p_{n+1} cannot exceed the smallest of these so that $p_{n+1} \leq p_1p_2\cdots p_n + 1$ for $n \geq 1$. Another way of saying the same thing is that

$$p_n \leq p_1p_2\cdots p_{n-1}+1 \qquad n \geq 2$$

With a slight modification of Euclid's reasoning, this inequality can be improved to give

$$p_n \leq p_1p_2\cdots p_{n-1}-1 \qquad n \geq 3$$

For instance, when $n = 5$, this tells us that

$$11 = p_5 \leq 2\cdot 3\cdot 5\cdot 7 - 1 = 209$$

We can see that the estimate is rather extravagant. A sharper limitation on the size of p_n is given by *Bonse's inequality*, which states that

$$p_n^2 < p_1 p_2 \cdots p_{n-1} \qquad n \geq 5$$

This inequality yields $p_5^2 < 210$, or $p_5 \leq 14$. A somewhat better size-estimate for p_5 comes from the inequality

$$p_{2n} \leq p_2 p_3 \cdots p_n - 2 \qquad n \geq 3$$

Here, we obtain

$$p_5 < p_6 \leq p_2 p_3 - 2 = 3 \cdot 5 - 2 = 13$$

To approximate the size of p_n from these formulas, it is necessary to know the values of $p_1, p_2, \ldots, p_{n-1}$. For a bound in which the preceding primes do not enter the picture, we have the following theorem.

Theorem 3.5. If p_n is the nth prime number, then $p_n \leq 2^{2^{n-1}}$.

Proof. Let us proceed by induction on n, the asserted inequality being clearly true when $n = 1$. As the hypothesis of the induction, we assume that $n > 1$ and that the result holds for all integers up to n. Then

$$\begin{aligned} p_{n+1} &\leq p_1 p_2 \cdots p_n + 1 \\ &\leq 2 \cdot 2^2 \cdots 2^{2^{n-1}} + 1 = 2^{1+2+2^2+\cdots+2^{n-1}} + 1 \end{aligned}$$

Recalling the identity $1 + 2 + 2^2 + \cdots + 2^{n-1} = 2^n - 1$, we obtain

$$p_{n+1} \leq 2^{2^n-1} + 1$$

However, $1 \leq 2^{2^n-1}$ for all n; whence

$$\begin{aligned} p_{n+1} &\leq 2^{2^n-1} + 2^{2^n-1} \\ &= 2 \cdot 2^{2^n-1} = 2^{2^n} \end{aligned}$$

completing the induction step, and the argument.

There is a corollary to Theorem 3.5 that is of interest.

Corollary. For $n \geq 1$, there are at least $n + 1$ primes less than 2^{2^n}.

Proof. From the theorem, we know that $p_1, p_2, \ldots, p_{n+1}$ are all less than 2^{2^n}.

We can do considerably better than is indicated by Theorem 3.5. In 1845, Joseph Bertrand conjectured that the prime numbers are well distributed in the sense that between $n \geq 2$ and $2n$ there is at least one prime. He was unable to establish his conjecture, but verified it for all $n \leq 3{,}000{,}000$. (One way of achieving this is to consider a sequence of primes 3, 5, 7, 13, 23, 43, 83, 163, 317, 631, 1259, 2503, 5003, 9973, 19937, 39869, 79699, 159389, ... each of which is less than twice the preceding.) Because it takes some real effort to substantiate this famous conjecture, let us content ourselves with saying that the first proof was carried out by the Russian mathematician P. L. Tchebycheff in 1852. Granting the result, it is not difficult to show that

$$p_n < 2^n \qquad n \geq 2$$

and as a direct consequence, $p_{n+1} < 2p_n$ for $n \geq 2$. In particular,

$$11 = p_5 < 2 \cdot p_4 = 14$$

To see that $p_n < 2^n$, we argue by induction on n. Clearly, $p_2 = 3 < 2^2$, so that the inequality is true here. Now assume that the inequality holds for an integer n, whence $p_n < 2^n$. Invoking Bertrand's conjecture, there exists a prime number p satisfying $2^n < p < 2^{n+1}$; that is, $p_n < p$. This immediately leads to the conclusion that $p_{n+1} \leq p < 2^{n+1}$, which completes the induction and the proof.

Primes of special form have been of perennial interest. Among these, the repunit primes are outstanding in their simplicity. A *repunit* is an integer written (in decimal notation) as a string of 1's, such as 11, 111, or 1111. Each such integer must have the form $(10^n - 1)/9$. We use the symbol R_n to denote the repunit consisting of n consecutive 1's. A peculiar feature of these numbers is the apparent scarcity of primes among them. So far, only R_2, R_{19}, R_{23}, R_{317}, R_{1031}, R_{49081}, R_{86453}, R_{109297}, and R_{270343} have been identified as primes (the last one in 2007). It is known that the only possible repunit primes R_n for all $n \leq 49000$ are the nine numbers just indicated. No conjecture has been made as to the existence of any others. For a repunit R_n to be prime, the subscript n must be a prime; that this is not a sufficient condition is shown by

$$R_5 = 11111 = 41 \cdot 271 \qquad R_7 = 1111111 = 239 \cdot 4649$$

PROBLEMS 3.2

1. Determine whether the integer 701 is prime by testing all primes $p \leq \sqrt{701}$ as possible divisors. Do the same for the integer 1009.
2. Employing the Sieve of Eratosthenes, obtain all the primes between 100 and 200.
3. Given that $p \nmid n$ for all primes $p \leq \sqrt[3]{n}$, show that $n > 1$ is either a prime or the product of two primes.
   [*Hint:* Assume to the contrary that n contains at least three prime factors.]
4. Establish the following facts:
   (a) $\sqrt{p}$ is irrational for any prime p.
   (b) If a is a positive integer and $\sqrt[n]{a}$ is rational, then $\sqrt[n]{a}$ must be an integer.
   (c) For $n \geq 2$, $\sqrt[n]{n}$ is irrational.
   [*Hint:* Use the fact that $2^n > n$.]
5. Show that any composite three-digit number must have a prime factor less than or equal to 31.
6. Fill in any missing details in this sketch of a proof of the infinitude of primes: Assume that there are only finitely many primes, say $p_1, p_2, \ldots, p_n$. Let A be the product of any r of these primes and put $B = p_1 p_2 \cdots p_n / A$. Then each p_k divides either A or B, but not both. Because $A + B > 1$, $A + B$ has a prime divisor different from any of the p_k, which is a contradiction.
7. Modify Euclid's proof that there are infinitely many primes by assuming the existence of a largest prime p and using the integer $N = p! + 1$ to arrive at a contradiction.
8. Give another proof of the infinitude of primes by assuming that there are only finitely many primes, say $p_1, p_2, \ldots, p_n$, and using the following integer to arrive at a contradiction:

$$N = p_2 p_3 \cdots p_n + p_1 p_3 \cdots p_n + \cdots + p_1 p_2 \cdots p_{n-1}$$

9. (a) Prove that if $n > 2$, then there exists a prime p satisfying $n < p < n!$.
[*Hint:* If $n! - 1$ is not prime, then it has a prime divisor p; and $p \leq n$ implies $p \mid n!$, leading to a contradiction.]
(b) For $n > 1$, show that every prime divisor of $n! + 1$ is an odd integer that is greater than n.

10. Let q_n be the smallest prime that is strictly greater than $P_n = p_1 p_2 \cdots p_n + 1$. It has been conjectured that the difference $q_n - (p_1 p_2 \cdots p_n)$ is always a prime. Confirm this for the first five values of n.

11. If p_n denotes the nth prime number, put $d_n = p_{n+1} - p_n$. An open question is whether the equation $d_n = d_{n+1}$ has infinitely many solutions. Give five solutions.

12. Assuming that p_n is the nth prime number, establish each of the following statements:
(a) $p_n > 2n - 1$ for $n \geq 5$.
(b) None of the integers $P_n = p_1 p_2 \cdots p_n + 1$ is a perfect square.
[*Hint:* Each P_n is of the form $4k + 3$ for $n > 1$.]
(c) The sum

$$\frac{1}{p_1} + \frac{1}{p_2} + \cdots + \frac{1}{p_n}$$

is never an integer.

13. For the repunits R_n, verify the assertions below:
(a) If $n \mid m$, then $R_n \mid R_m$.
[*Hint:* If $m = kn$, consider the identity

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

(b) If $d \mid R_n$ and $d \mid R_m$, then $d \mid R_{n+m}$.
[*Hint:* Show that $R_{m+n} = R_n 10^m + R_m$.]
(c) If $\gcd(n, m) = 1$, then $\gcd(R_n, R_m) = 1$.

14. Use the previous problem to obtain the prime factors of the repunit R_{10}.

3.3 THE GOLDBACH CONJECTURE

Although there is an infinitude of primes, their distribution within the positive integers is most mystifying. Repeatedly in their distribution we find hints or, as it were, shadows of a pattern; yet an actual pattern amenable to precise description remains elusive. The difference between consecutive primes can be small, as with the pairs 11 and 13, 17 and 19, or for that matter 1000000000061 and 1000000000063. At the same time there exist arbitrarily long intervals in the sequence of integers that are totally devoid of any primes.

It is an unanswered question whether there are infinitely many pairs of *twin primes*; that is, pairs of successive odd integers p and $p + 2$ that are both primes. Numerical evidence leads us to suspect an affirmative conclusion. Electronic computers have discovered 152891 pairs of twin primes less than 30000000 and 20 pairs between 10^{12} and 10^{12}+ 10000, which hints at their growing scarcity as the positive integers increase in magnitude. Many examples of immense twins are known. The largest twins to date, each 100355 digits long,

$$65516468355 \cdot 2^{333333} \pm 1$$

were discovered in 2009.

Consecutive primes not only can be close together, but also can be far apart; that is, arbitrarily large gaps can occur between consecutive primes. Stated precisely: Given any positive integer n, there exist n consecutive integers, all of which are composite. To prove this, we simply need to consider the integers

$$(n + 1)! + 2, (n + 1)! + 3, \ldots, (n + 1)! + (n + 1)$$

where $(n + 1)! = (n + 1) \cdot n \cdots 3 \cdot 2 \cdot 1$. Clearly there are n integers listed, and they are consecutive. What is important is that each integer is composite. Indeed, $(n + 1)! + 2$ is divisible by 2, $(n + 1)! + 3$ is divisible by 3, and so on.

For instance, if a sequence of four consecutive composite integers is desired, then the previous argument produces 122, 123, 124, and 125:

$$\begin{aligned}
5! + 2 &= 122 = 2 \cdot 61\\
5! + 3 &= 123 = 3 \cdot 41\\
5! + 4 &= 124 = 4 \cdot 31\\
5! + 5 &= 125 = 5 \cdot 25
\end{aligned}$$

Of course, we can find other sets of four consecutive composites, such as 24, 25, 26, 27 or 32, 33, 34, 35.

As this example suggests, our procedure for constructing gaps between two consecutive primes gives a gross overestimate of where they occur among the integers. The first occurrences of prime gaps of specific lengths, where all the intervening integers are composite, have been the subject of computer searches. For instance, there is a gap of length 778 (that is, $p_{n+1} - p_n = 778$) following the prime 42842283925351. No gap of this size exists between two smaller primes. The largest effectively calculated gap between consecutive prime numbers has length 1442, with a string of 1441 composites immediately after the prime

$$804212830686677669$$

Interestingly, computer researchers have not identified gaps of every possible width up to 1442. The smallest missing gap size is 796. The conjecture is that there is a prime gap (a string of $2k - 1$ consecutive composites between two primes) for every even integer $2k$.

This brings us to another unsolved problem concerning the primes, the Goldbach conjecture. In a letter to Leonhard Euler in the year 1742, Christian Goldbach hazarded the guess that every even integer is the sum of two numbers that are either primes or 1. A somewhat more general formulation is that every even integer greater than 4 can be written as a sum of two odd prime numbers. This is easy to confirm for the first few even integers:

$$\begin{aligned}
2 &= 1 + 1\\
4 &= 2 + 2 = 1 + 3\\
6 &= 3 + 3 = 1 + 5\\
8 &= 3 + 5 = 1 + 7\\
10 &= 3 + 7 = 5 + 5\\
12 &= 5 + 7 = 1 + 11\\
14 &= 3 + 11 = 7 + 7 = 1 + 13\\
16 &= 3 + 13 = 5 + 11\\
18 &= 5 + 13 = 7 + 11 = 1 + 17\\
20 &= 3 + 17 = 7 + 13 = 1 + 19\\
22 &= 3 + 19 = 5 + 17 = 11 + 11\\
24 &= 5 + 19 = 7 + 17 = 11 + 13 = 1 + 23\\
26 &= 3 + 23 = 7 + 19 = 13 + 13\\
28 &= 5 + 23 = 11 + 17\\
30 &= 7 + 23 = 11 + 19 = 13 + 17 = 1 + 29
\end{aligned}$$

Although it seems that Euler never tried to prove the result, upon writing to Goldbach at a later date, Euler countered with a conjecture of his own: Any even integer (≥ 6) of the form $4n + 2$ is a sum of two numbers each being either a prime of the form $4n + 1$ or 1.

The numerical data suggesting the truth of Goldbach's conjecture are overwhelming. It has been verified by computers for all even integers less than $4 \cdot 10^{14}$. As the integers become larger, the number of different ways in which $2n$ can be expressed as the sum of two primes increases. For example, there are 291400 such representations for the even integer 100000000. Although this supports the feeling that Goldbach was correct in his conjecture, it is far from a mathematical proof, and all attempts to obtain a proof have been completely unsuccessful. One of the most famous number theorists of the last century, G. H. Hardy, in his address to the Mathematical Society of Copenhagen in 1921, stated that the Goldbach conjecture appeared "probably as difficult as any of the unsolved problems in mathematics." It is currently known that every even integer is the sum of six or fewer primes.

We remark that if the conjecture of Goldbach is true, then each odd number larger than 7 must be the sum of three odd primes. To see this, take n to be an odd integer greater than 7, so that $n - 3$ is even and greater than 4; if $n - 3$ could be expressed as the sum of two odd primes, then n would be the sum of three.

The first real progress on the conjecture in nearly 200 years was made by Hardy and Littlewood in 1922. On the basis of a certain unproved hypothesis, the so-called generalized Riemann hypothesis, they showed that every sufficiently large odd number is the sum of three odd primes. In 1937, the Russian mathematician I. M. Vinogradov was able to remove the dependence on the generalized Riemann hypothesis, thereby giving an unconditional proof of this result; that is to say, he established that all odd integers greater than some effectively computable n_0 can be written as the sum of three odd primes.

$$n = p_1 + p_2 + p_3 \qquad (n \text{ odd, } n \text{ sufficiently large})$$

Vinogradov was unable to decide how large n_0 should be, but Borozdkin (1956) proved that $n_0 < 3^{3^{15}}$. In 2002, the bound on n_0 was reduced to 10^{1346}. It follows immediately that every even integer from some point on is the sum of either two or four primes. Thus, it is enough to answer the question for every odd integer n in the range $9 \leq n \leq n_0$, which, for a given integer, becomes a matter of tedious computation (unfortunately, n_0 is so large that this exceeds the capabilities of the most modern electronic computers).

Because of the strong evidence in favor of the famous Goldbach conjecture, we readily become convinced that it is true. Nevertheless, it might be false. Vinogradov showed that if $A(x)$ is the number of even integers $n \leq x$ that are not the sum of two primes, then

$$\lim_{x \to \infty} A(x)/x = 0$$

This allows us to say that "almost all" even integers satisfy the conjecture. As Edmund Landau so aptly put it, "The Goldbach conjecture is false for at most 0% of all even integers; this *at most* 0% does not exclude, of course, the possibility that there are infinitely many exceptions."

Having digressed somewhat, let us observe that according to the Division Algorithm, every positive integer can be written uniquely in one of the forms

$$4n \qquad 4n+1 \qquad 4n+2 \qquad 4n+3$$

for some suitable $n \geq 0$. Clearly, the integers $4n$ and $4n + 2 = 2(2n + 1)$ are both even. Thus, all odd integers fall into two progressions: one containing integers of the form $4n + 1$, and the other containing integers of the form $4n + 3$.

The question arises as to how these two types of primes are distributed within the set of positive integers. Let us display the first few odd prime numbers in consecutive order, putting the $4n + 3$ primes in the top row and the $4n + 1$ primes under them:

3	7	11	19	23	31	43	47	59	67	71	79	83
5	13	17	29	37	41	53	61	73	89			

At this point, one might have the general impression that primes of the form $4n + 3$ are more abundant than are those of the form $4n + 1$. To obtain more precise information, we require the help of the function $\pi_{a,b}(x)$, which counts the number of primes of the form $p = an + b$ not exceeding x. Our small table, for instance, indicates that $\pi_{4,1}(89) = 10$ and $\pi_{4,3}(89) = 13$.

In a famous letter written in 1853, Tchebycheff remarked that $\pi_{4,1}(x) \leq \pi_{4,3}(x)$ for small values of x. He also implied that he had a proof that the inequality always held. In 1914, J. E. Littlewood showed that the inequality fails infinitely often, but his method gave no indication of the value of x for which this first happens. It turned out to be quite difficult to find. Not until 1957 did a computer search reveal that $x = 26861$ is the smallest prime for which $\pi_{4,1}(x) > \pi_{4,3}(x)$; here, $\pi_{4,1}(x) = 1473$

and $\pi_{4,3}(x) = 1472$. This is an isolated situation, because the next prime at which a reversal occurs is $x = 616{,}841$. Remarkably, $\pi_{4,1}(x) > \pi_{4,3}(x)$ for the 410 million successive integers x lying between 18540000000 and 18950000000.

The behavior of primes of the form $3n \pm 1$ provided more of a computational challenge: the inequality $\pi_{3,1}(x) \leq \pi_{3,2}(x)$ holds for all x until one reaches $x = 608981813029$.

This furnishes a pleasant opportunity for a repeat performance of Euclid's method for proving the existence of an infinitude of primes. A slight modification of his argument reveals that there is an infinite number of primes of the form $4n + 3$. We approach the proof through a simple lemma.

Lemma. The product of two or more integers of the form $4n + 1$ is of the same form.

Proof. It is sufficient to consider the product of just two integers. Let us take $k = 4n + 1$ and $k' = 4m + 1$. Multiplying these together, we obtain

$$\begin{aligned} kk' &= (4n + 1)(4m + 1) \\ &= 16nm + 4n + 4m + 1 = 4(4nm + n + m) + 1 \end{aligned}$$

which is of the desired form.

This paves the way for Theorem 3.6.

Theorem 3.6. There are an infinite number of primes of the form $4n + 3$.

Proof. In anticipation of a contradiction, let us assume that there exist only finitely many primes of the form $4n + 3$; call them $q_1, q_2, \ldots, q_s$. Consider the positive integer

$$N = 4q_1q_2 \cdots q_s - 1 = 4(q_1q_2 \cdots q_s - 1) + 3$$

and let $N = r_1r_2 \cdots r_t$ be its prime factorization. Because N is an odd integer, we have $r_k \neq 2$ for all k, so that each r_k is either of the form $4n + 1$ or $4n + 3$. By the lemma, the product of any number of primes of the form $4n + 1$ is again an integer of this type. For N to take the form $4n + 3$, as it clearly does, N must contain at least one prime factor r_i of the form $4n + 3$. But r_i cannot be found among the listing $q_1, q_2, \ldots, q_s$, for this would lead to the contradiction that $r_i \mid 1$. The only possible conclusion is that there are infinitely many primes of the form $4n + 3$.

Having just seen that there are infinitely many primes of the form $4n + 3$, we might reasonably ask: Is the number of primes of the form $4n + 1$ also infinite? This answer is likewise in the affirmative, but a demonstration must await the development of the necessary mathematical machinery. Both these results are special cases of a remarkable theorem by P. G. L. Dirichlet on primes in arithmetic progressions, established in 1837. The proof is much too difficult for inclusion here, so that we must content ourselves with the mere statement.

Theorem 3.7 Dirichlet. If a and b are relatively prime positive integers, then the arithmetic progression

$$a, a + b, a + 2b, a + 3b, \ldots$$

contains infinitely many primes.

Dirichlet's theorem tells us, for instance, that there are infinitely many prime numbers ending in 999, such as 1999, 100999, 1000999, ... for these appear in the arithmetic progression determined by $1000n + 999$, where $\gcd(1000, 999) = 1$.

There is no arithmetic progression $a, a + b, a + 2b, \ldots$ that consists solely of prime numbers. To see this, suppose that $a + nb = p$, where p is a prime. If we put $n_k = n + kp$ for $k = 1, 2, 3, \ldots$ then the n_kth term in the progression is

$$a + n_k b = a + (n + kp)b = (a + nb) + kpb = p + kpb$$

Because each term on the right-hand side is divisible by p, so is $a + n_k b$. In other words, the progression must contain infinitely many composite numbers.

It was proved in 2008 that there are finite but arbitrarily long arithmetic progressions consisting only of prime numbers (not necessarily consecutive primes). The longest progression found to date is composed of the 23 primes:

$$56211383760397 + 44546738095860n \quad 0 \leq n \leq 22$$

The prime factorization of the common difference between the terms is

$$2^2 \cdot 3 \cdot 5 \cdot 7 \cdot 11 \cdot 13 \cdot 17 \cdot 19 \cdot 23 \cdot 99839$$

which is divisible by 9699690, the product of the primes less than 23. This takes place according to Theorem 3.8.

Theorem 3.8. If all the $n > 2$ terms of the arithmetic progression

$$p, p + d, p + 2d, \ldots, p + (n - 1)d$$

are prime numbers, then the common difference d is divisible by every prime $q < n$.

Proof. Consider a prime number $q < n$ and assume to the contrary that $q \nmid d$. We claim that the first q terms of the progression

$$p, p + d, p + 2d, \ldots, p + (q - 1)d \tag{1}$$

will leave different remainders when divided by q. Otherwise there exist integers j and k, with $0 \leq j < k \leq q - 1$, such that the numbers $p + jd$ and $p + kd$ yield the same remainder upon division by q. Then q divides their difference $(k - j)d$. But $\gcd(q, d) = 1$, and so Euclid's lemma leads to $q \mid k - j$, which is nonsense in light of the inequality $k - j \leq q - 1$.

Because the q different remainders produced from Eq. (1) are drawn from the q integers $0, 1, \ldots, q - 1$, one of these remainders must be zero. This means that $q \mid p + td$ for some t satisfying $0 \leq t \leq q - 1$. Because of the inequality $q < n \leq p \leq p + td$, we are forced to conclude that $p + td$ is composite. (If p were less than n, one of the terms of the progression would be $p + pd = p(1 + d)$.) With this contradiction, the proof that $q \mid d$ is complete.

It has been conjectured that there exist arithmetic progressions of finite (but otherwise arbitrary) length, composed of consecutive prime numbers. Examples of such progressions consisting of three and four primes, respectively, are 47, 53, 59, and 251, 257, 263, 269.

Most recently a sequence of 10 consecutive primes was discovered in which each term exceeds its predecessor by just 210; the smallest of these primes has 93 digits.

Finding an arithmetic progression consisting of 11 consecutive primes is likely to be out of reach for some time. Absent the restriction that the primes involved be consecutive, strings of 11-term arithmetic progressions are easily located. One such is

$$110437 + 13860n \qquad 0 \leq n \leq 10$$

In the interest of completeness, we might mention another famous problem that, so far, has resisted the most determined attack. For centuries, mathematicians have sought a simple formula that would yield every prime number or, failing this, a formula that would produce nothing but primes. At first glance, the request seems modest enough: find a function $f(n)$ whose domain is, say, the nonnegative integers and whose range is some infinite subset of the set of all primes. It was widely believed years ago that the quadratic polynomial

$$f(n) = n^2 + n + 41$$

assumed only prime values. This was shown to be false by Euler, in 1772. As evidenced by the following table, the claim is a correct one for $n = 0, 1, 2, \ldots, 39$.

n	$f(n)$	n	$f(n)$	n	$f(n)$
0	41	14	251	28	853
1	43	15	281	29	911
2	47	16	313	30	971
3	53	17	347	31	1033
4	61	18	383	32	1097
5	71	19	421	33	1163
6	83	20	461	34	1231
7	97	21	503	35	1301
8	113	22	547	36	1373
9	131	23	593	37	1447
10	151	24	641	38	1523
11	173	25	691	39	1601
12	197	26	743		
13	223	27	797		

However, this provocative conjecture is shattered in the cases $n = 40$ and $n = 41$, where there is a factor of 41:

$$f(40) = 40 \cdot 41 + 41 = 41^2$$

and

$$f(41) = 41 \cdot 42 + 41 = 41 \cdot 43$$

The next value $f(42) = 1847$ turns out to be prime once again. In fact, for the first 100 integer values of n, the so-called Euler polynomial represents 86 primes. Although it starts off very well in the production of primes, there are other quadratics such as

$$g(n) = n^2 + n + 27941$$

that begin to best $f(n)$ as the values of n become larger. For example, $g(n)$ is prime for 286129 values of $0 \leq n \leq 10^6$, whereas its famous rival yields 261081 primes in this range.

It has been shown that no polynomial of the form $n^2 + n + q$, with q a prime, can do better than the Euler polynomial in giving primes for successive values of n. Indeed, until fairly recently no other quadratic polynomial of any kind was known to produce more than 40 successive prime values. The polynomial

$$h(n) = 103n^2 - 3945n + 34381$$

found in 1988, produces 43 distinct prime values for $n = 0, 1, 2, \ldots, 42$. The current record holder in this regard

$$k(n) = 36n^2 - 810n + 2753$$

does slightly better by giving a string of 45 prime values.

The failure of the previous functions to be prime-producing is no accident, for it is easy to prove that there is no nonconstant polynomial $f(n)$ with integral coefficients that takes on just prime values for integral $n \geq 0$. We assume that such a polynomial $f(n)$ actually does exist and argue until a contradiction is reached. Let

$$f(n) = a_k n^k + a_{k-1} n^{k-1} + \cdots + a_2 n^2 + a_1 n + a_0$$

where all the coefficients $a_0, a_1, \ldots, a_k$ are integers, and $a_k \neq 0$. For a fixed value of (n_0), $p = f(n_0)$ is a prime number. Now, for any integer t, we consider the following expression:

$$\begin{aligned} f(n_0 + tp) &= a_k(n_0 + tp)^k + \cdots + a_1(n_0 + tp) + a_0 \\ &= (a_k n_0^k + \cdots + a_1 n_0 + a_0) + pQ(t) \\ &= f(n_0) + pQ(t) \\ &= p + pQ(t) = p(1 + Q(t)) \end{aligned}$$

where $Q(t)$ is a polynomial in t having integral coefficients. Our reasoning shows that $p \mid f(n_0 + tp)$; hence, from our own assumption that $f(n)$ takes on only prime values, $f(n_0 + tp) = p$ for any integer t. Because a polynomial of degree k cannot assume the same value more than k times, we have obtained the required contradiction.

Recent years have seen a measure of success in the search for prime-producing functions. W. H. Mills proved (1947) that there exists a positive real number r such that the expression $f(n) = [r^{3^n}]$ is prime for $n = 1, 2, 3, \ldots$ (the brackets indicate the greatest integer function). Needless to say, this is strictly an existence theorem and nothing is known about the actual value of r. Mills's function does not produce all the primes.

There are several celebrated, still unresolved, conjectures about primes. One posed by G. H. Hardy and J. E. Littlewood in 1922 asks whether there are infinitely many primes that can be represented in the form $n^2 + 1$. The closest thing to an answer, so far, came in 1978 when it was proved that there are infinitely many values of n for which $n^2 + 1$ is either a prime or the product of just two primes. One can start to see this for the smallest values

$$\begin{array}{lll} 2^2+1=5 & 5^2+1=2\cdot 13 & 9^2+1=2\cdot 41 \\ 3^2+1=2\cdot 5 & 6^2+1=37 & 10^2+1=101 \\ 4^2+1=17 & 8^2+1=5\cdot 31 & \end{array}$$

PROBLEMS 3.3

1. Verify that the integers 1949 and 1951 are twin primes.
2. (a) If 1 is added to a product of twin primes, prove that a perfect square is always obtained.
   (b) Show that the sum of twin primes p and $p+2$ is divisible by 12, provided that $p > 3$.
3. Find all pairs of primes p and q satisfying $p - q = 3$.
4. Sylvester (1896) rephrased the Goldbach conjecture: Every even integer $2n$ greater than 4 is the sum of two primes, one larger than $n/2$ and the other less than $3n/2$. Verify this version of the conjecture for all even integers between 6 and 76.
5. In 1752, Goldbach submitted the following conjecture to Euler: Every odd integer can be written in the form $p + 2a^2$, where p is either a prime or 1 and $a \geq 0$. Show that the integer 5777 refutes this conjecture.
6. Prove that the Goldbach conjecture that every even integer greater than 2 is the sum of two primes is equivalent to the statement that every integer greater than 5 is the sum of three primes.
   [*Hint:* If $2n - 2 = p_1 + p_2$, then $2n = p_1 + p_2 + 2$ and $2n + 1 = p_1 + p_2 + 3$.]
7. A conjecture of Lagrange (1775) asserts that every odd integer greater than 5 can be written as a sum $p_1 + 2p_2$, where p_1, p_2 are both primes. Confirm this for all odd integers through 75.
8. Given a positive integer n, it can be shown that there exists an even integer a that is representable as the sum of two odd primes in n different ways. Confirm that the integers 60, 78, and 84 can be written as the sum of two primes in six, seven, and eight ways, respectively.
9. (a) For $n > 3$, show that the integers $n, n+2, n+4$ cannot all be prime.
   (b) Three integers p, $p+2$, $p+6$, which are all prime, are called a *prime-triplet*. Find five sets of prime-triplets.
10. Establish that the sequence

$$(n+1)! - 2, (n+1)! - 3, \ldots, (n+1)! - (n+1)$$

   produces n consecutive composite integers for $n > 2$.
11. Find the smallest positive integer n for which the function $f(n) = n^2 + n + 17$ is composite. Do the same for the functions $g(n) = n^2 + 21n + 1$ and $h(n) = 3n^2 + 3n + 23$.
12. Let p_n denote the nth prime number. For $n \geq 3$, prove that $p_{n+3}^2 < p_n p_{n+1} p_{n+2}$.
    [*Hint:* Note that $p_{n+3}^2 < 4p_{n+2}^2 < 8p_{n+1}p_{n+2}$.]
13. Apply the same method of proof as in Theorem 3.6 to show that there are infinitely many primes of the form $6n + 5$.
14. Find a prime divisor of the integer $N = 4(3\cdot 7\cdot 11) - 1$ of the form $4n+3$. Do the same for $N = 4(3\cdot 7\cdot 11\cdot 15) - 1$.
15. Another unanswered question is whether there exists an infinite number of sets of five consecutive odd integers of which four are primes. Find five such sets of integers.
16. Let the sequence of primes, with 1 adjoined, be denoted by $p_0 = 1$, $p_1 = 2$, $p_2 = 3$, $p_3 = 5, \ldots$. For each $n \geq 1$, it is known that there exists a suitable choice of coefficients

$\epsilon_k = \pm 1$ such that

$$p_{2n} = p_{2n-1} + \sum_{k=0}^{2n-2} \epsilon_k p_k \qquad p_{2n+1} = 2p_{2n} + \sum_{k=0}^{2n-1} \epsilon_k p_k$$

To illustrate:

$$13 = 1 + 2 - 3 - 5 + 7 + 11$$

and

$$17 = 1 + 2 - 3 - 5 + 7 - 11 + 2 \cdot 13$$

Determine similar representations for the primes 23, 29, 31, and 37.

17. In 1848, de Polignac claimed that every odd integer is the sum of a prime and a power of 2. For example, $55 = 47 + 2^3 = 23 + 2^5$. Show that the integers 509 and 877 discredit this claim.

18. (a) If p is a prime and $p \nmid b$, prove that in the arithmetic progression

$$a, a + b, a + 2b, a + 3b, \ldots$$

every pth term is divisible by p.
[*Hint:* Because $\gcd(p, b) = 1$, there exist integers r and s satisfying $pr + bs = 1$. Put $n_k = kp - as$ for $k = 1, 2, \ldots$ and show that $p \mid (a + n_k b)$.]

(b) From part (a), conclude that if b is an odd integer, then every other term in the indicated progression is even.

19. In 1950, it was proved that any integer $n > 9$ can be written as a sum of distinct odd primes. Express the integers 25, 69, 81, and 125 in this fashion.

20. If p and $p^2 + 8$ are both prime numbers, prove that $p^3 + 4$ is also prime.

21. (a) For any integer $k > 0$, establish that the arithmetic progression

$$a + b, a + 2b, a + 3b, \ldots$$

where $\gcd(a, b) = 1$, contains k consecutive terms that are composite.
[*Hint:* Put $n = (a + b)(a + 2b) \cdots (a + kb)$ and consider the k terms $a + (n + 1)b$, $a + (n + 2)b, \ldots, a + (n + k)b$.]

(b) Find five consecutive composite terms in the arithmetic progression

$$6, 11, 16, 21, 26, 31, 36, \ldots$$

22. Show that 13 is the largest prime that can divide two successive integers of the form $n^2 + 3$.

23. (a) The arithmetic mean of the twin primes 5 and 7 is the triangular number 6. Are there any other twin primes with a triangular mean?

(b) The arithmetic mean of the twin primes 3 and 5 is the perfect square 4. Are there any other twin primes with a square mean?

24. Determine all twin primes p and $q = p + 2$ for which $pq - 2$ is also prime.

25. Let p_n denote the nth prime. For $n > 3$, show that

$$p_n < p_1 + p_2 + \cdots + p_{n-1}$$

[*Hint:* Use induction and the Bertrand conjecture.]

26. Verify the following:

(a) There exist infinitely many primes ending in 33, such as 233, 433, 733, 1033, ...
[*Hint:* Apply Dirichlet's theorem.]

(b) There exist infinitely many primes that do not belong to any pair of twin primes.
[*Hint:* Consider the arithmetic progression $21k + 5$ for $k = 1, 2, \ldots$.]

(c) There exists a prime ending in as many consecutive 1's as desired.
[*Hint:* To obtain a prime ending in n consecutive 1's, consider the arithmetic progression $10^n k + R_n$ for $k = 1, 2, \ldots$.]

(d) There exist infinitely many primes that contain but do not end in the block of digits 123456789.
[*Hint:* Consider the arithmetic progression $10^{11}k + 1234567891$ for $k = 1, 2, \ldots$.]

27. Prove that for every $n \geq 2$ there exists a prime p with $p \leq n < 2p$.
[*Hint:* In the case where $n = 2k + 1$, then by the Bertrand conjecture there exists a prime p such that $k < p < 2k$.]

28. (a) If $n > 1$, show that $n!$ is never a perfect square.

(b) Find the values of $n \geq 1$ for which

$$n! + (n + 1)! + (n + 2)!$$

is a perfect square.
[*Hint:* Note that $n! + (n + 1)! + (n + 2)! = n!(n + 2)^2$.]