Foundations of Arithmetic

Notation

We shall denote the sum and product of numbers in the usual notation as

$$a_1 + a_2 + a_3 + \cdots + a_k = \sum_{i=1}^{k} a_i\,, \quad a_1 a_2 a_3 \ldots a_k = \prod_{i=1}^{k} a_i$$

The notation $a|b$ means a divides b, i.e. $ac = b$ where c is an integer, and $a \nmid b$ means a does not divide b. If $ac = b$ $(|b| > 1)$ implies $a = \pm 1$ or $a = \pm b$, then b is a *prime number*. A number b $(|b| > 1)$ that is not prime is said to be *composite*. There are infinitely many primes, for if there were only a finite number, $p_1, p_1 \ \ldots \ p_n$ say, then the number $p_1 p_2$ ... p_n+1 is not divisible by any of them and is therefore a new prime itself or is divisible by some other new prime, which contradicts the assumption that there are only n primes.

Let

$$\max(a,b) = \begin{cases} a & (a > b) \\ b & (b > a) \end{cases}, \quad \min(a,b) = \begin{cases} b & (a > b) \\ a & (b > a) \end{cases}, \quad \max(a,a) = \min(a,a) = a.$$

It is obvious that

$$\max(a,b) + \min(a,b) = a + b. \tag{1}$$

The generalization of this notation to three numbers a, b, c is straightforward, the only difference being that one of the three will be a middle number that is neither the maximum nor the minimum (although it could be equal to one of them). Consider the three pairs $(a,b), (b,c)$ and (c,a). Only one of them involves the maximum and the middle number. The minimum of this pair will therefore be the middle number. The other two pairs must involve the minimum of the three numbers so that the minimum of both pairs will be $\min(a,b,c)$ and therefore their sum will be $2 \times \min(a,b,c)$. It follows that

$$\min(a,b) + \min(b,c) + \min(c,a) + \max(a,b,c) - \min(a,b,c) = a + b + c$$

which is the equation that corresponds to (1). Note that if $c = b$ then $\max(a,b,c) = \max(a,b)$, $\min(a,b,c) = \min(a,b)$, $\min(c,a) = \min(a,b)$ and $\min(b,b) = b$. Substitution in the equation above reduces it to (1), as required.

The fundamental theorem of arithmetic

The Fundamental Theorem of Arithmetic states that any number $m \geq 1$ is expressed uniquely as a product of powers of prime numbers. Denote the i^{th} prime number by p_i, i.e. $p_1 = 2$, $p_2 = 3$, $p_3 = 5$, ... , $p_8 = 19$, etc. Then

$$m = \prod_{i=1}^{\infty} p_i^{j_i} \quad (j_i \geq 0). \tag{2}$$

If p_k, say, is not a prime factor of m then $j_k = 0$ so that the factor $p_k^{j_k} = 1$. If p_M is the largest prime factor of m then $j_i = 0$ for $i > M$ so that all the factors involving primes greater than p_M reduce to 1. If $m = 1$ then $j_i = 0$ for $i \geq 1$

The representation (2) can be proved by induction. Assume it is true for $1 \leq m \leq n-1$. When $m = n$, n can be either prime, in which case the representation (2) is trivially true, or it is composite and can therefore be expressed as $n = ab$, where $1 < a \leq n-1$ and $1 < b \leq n-1$. But by the induction hypothesis both a and b can be expressed as a product of primes as in (2) so that the product ab can also be expressed as a product of primes. It follows that (2) is true for $m = n = ab$ and by induction the result is proved for all m.

It remains to be shown that (2) is unique. Suppose there are two different representations of m expressed in the form of (2), i.e.

$$m = \prod_{i=1}^{\infty} p_i^{j_i} = \prod_{i=1}^{\infty} p_i^{k_i} \quad (j_i \geq 0, \ \ k_i \geq 0). \tag{3}$$

Consider a prime factor p_s for which $\max(j_s, k_s) \neq 0$ and $j_s \neq k_s$ and assume for the sake of argument that $j_s > k_s$. Then (3) can be written as

$$p_s^{j_s - k_s} \prod_{i=1}^{s-1} p_i^{j_i} \prod_{i=s+1}^{\infty} p_i^{j_i} = \prod_{i=1}^{s-1} p_i^{k_i} \prod_{i=s+1}^{\infty} p_i^{k_i} \quad (j_i \geq 0, \ \ k_i \geq 0).$$

(Note if $k_s > j_s$, a factor $p_s^{k_s - j_s}$ would be taken out of the right-hand side instead.) If $j_s \neq k_s$ the left-hand side of this equation is divisible by p_s but the right-hand side is not, which is impossible. Thus $j_s = k_s$ and since this applies to all prime factors p_i whose exponents are not already equal, we deduce that the two products in (3) are identical. Therefore, the representation of a number m as a product of primes is unique.

Let us now consider a couple of consequences of the Fundamental Theorem.

(i) ***The highest common factor and lowest common multiple of two and three numbers***

By analogy with (2), another number $n \geq 1$ can be represented as

$$n = \prod_{i=1}^{\infty} p_i^{k_i} \quad (k_i \geq 0)$$

where $k_i = 0$ for $i > N$ when p_N is the largest prime factor of n. The Highest Common Factor (HCF) of two numbers m and n is the largest divisor of both numbers (which is why it is often called the Greatest Common Divisor or GCD). Clearly the HCF must include all the common prime factors of both numbers. If p_i is a common factor raised to the powers j_i and k_i

respectively, then the largest factor common to both m and n is $p_i^{\min(j_i,k_i)}$. Note that if p_i is *not* a common factor but occurs only in the representation of m (say) then it is necessary that $k_i = 0$ in order to exclude it from the representation of n, as required. In this case $\min(j_i, k_i) = 0$ so that p_i is omitted from the HCF as well, as indeed it should be. Thus all cases are included in the formal definition

$$\text{hcf}(m,n) = \prod_{i=1}^{\infty} p_i^{\min(j_i,k_i)}$$

The Lowest Common Multiple (LCM) is the smallest number that is divisible by both m and n. This time we must choose p_i raised to the greater of the two powers j_i and k_i for both m and n to be divisors of the LCM. Thus we define

$$\text{lcm}(m,n) = \prod_{i=1}^{\infty} p_i^{\max(j_i,k_i)}$$

Using (1) we obtain

$$\prod_{i=1}^{\infty} p_i^{\min(j_i,k_i)} \prod_{i=1}^{\infty} p_i^{\max(j_i,k_i)} = \prod_{i=1}^{\infty} p_i^{\max(j_i,k_i)+\min(j_i,k_i)} = \prod_{i=1}^{\infty} p_i^{j_i+\,k_i} = \prod_{i=1}^{\infty} p_i^{j_i} \prod_{i=1}^{\infty} p_i^{k_i}$$

Hence

$$\text{hcf}(m,n) \cdot \text{lcm}(m,n) = mn.$$

Thus once the HCF of two numbers is known, the LCM is easily found.

With $q \geq 1$ defined by

$$q = \prod_{i=1}^{\infty} p_i^{l_i} \quad (l_i \geq 0)$$

the definitions of the HCF and LCM of three numbers m, n and q become

$$\text{hcf}(m,n,q) = \prod_{i=1}^{\infty} p_i^{\min(j_i,k_i,l_i)}, \qquad \text{lcm}(m,n,q) = \prod_{i=1}^{\infty} p_i^{\max(j_i,k_i,l_i)}$$

Let $\text{hcf}(m,n) = h$ and $\min(j_i, k_i) = s_i$. Since $\min(j_i, k_i, l_i) = \min(s_i, l_i)$, it is clear that

$$\text{hcf}(h,q) = \prod_{i=1}^{\infty} p_i^{\min(s_i,l_i)} = \prod_{i=1}^{\infty} p_i^{\min(j_i,k_i,l_i)} = \text{hcf}(m,n,q)$$

which demonstrates the fairly obvious fact that the HCF of three numbers is the HCF of the pair comprising one number and the HCF of the other two, i.e. $\text{hcf}(m,n,q) = \text{hcf}[\text{hcf}(m,n), q]$.

The relation between the HCF and LCM of three numbers is more complicated. Following the same approach as before, but using the modified formula for the maximum of three numbers instead of (1), we obtain

$$\prod_{i=1}^{\infty} p_i^{\max(j_i,k_i,l_i)} = \prod_{i=1}^{\infty} p_i^{j_i+k_i+l_i+\min(j_i,k_i,l_i)-\min(j_i,k_i)-\min(k_i,l_i)-\min(l_i,j_i)}$$

which gives the relation

$$\mathrm{lcm}(m,n,q) = \frac{mnq\ \mathrm{hcf}(m,n,q)}{\mathrm{hcf}(m,n)\mathrm{hcf}(n,q)\mathrm{hcf}(q,m)}$$

If $q = n$, then $\mathrm{lcm}(m,n,q) = \mathrm{lcm}(m,n)$, $\mathrm{hcf}(m,n,q)/\mathrm{hcf}(m,n) = 1$, $q/\mathrm{hcf}(n,q) = 1$ and $\mathrm{hcf}(q,m) = \mathrm{hcf}(m,n)$ which substituted above give $\mathrm{lcm}(m,n) = mn/\mathrm{hcf}(m,n)$, in agreement with the relation for two numbers.

There appears to be little point in deriving a connection between the LCM and HCF for more than three numbers as it will become increasingly complicated and of little practical help.

(ii) ***Euclid's algorithm and numerical examples***

A familiar way of calculating the HCF of two numbers dates back to Euclid. Let $h = \mathrm{hcf}(m,n)$ and suppose $m \geq n$. Then

$$m = c_1 n + r_1 \quad (0 \leq r_1 < n). \tag{4}$$

Here r_1 is the remainder left after dividing m by n. If $n|m$ then $r_1 = 0$ and the HCF is simply n itself. Otherwise r_1 must be smaller than n because c_1 is the maximum number of times n goes into m. Since $h|m$ and $h|n$ it is obvious from (4) that $h|r_1$ as well, and because $n > r_1$ we may therefore write by analogy with (4)

$$n = c_2 r_1 + r_2 \quad (0 \leq r_2 < r_1). \tag{5}$$

Again we have $h|r_2$ because it also divides both n and r_1 in equation (5). The procedure continues in this way, the next step yielding

$$r_1 = c_3 r_2 + r_3 \quad (0 \leq r_3 < r_2)$$

where $h|r_3$ and so on until we reach the $(k-1)^{\text{th}}$ and k^{th} steps

$$r_{k-2} = c_k r_{k-1} + r_k \quad (0 \leq r_k < r_{k-1}) \tag{6}$$

$$r_{k-1} = c_{k+1} r_k + r_{k+1} \quad (0 \leq r_{k+1} < r_k).$$

The sequence of positive integers $r_1 > r_2 > \cdots > r_{k-1} > r_k$ is decreasing so must eventually terminate in 0. Let $r_{k+1} = 0$ so that the last equation above reduces to

$$r_{k-1} = c_{k+1} r_k \quad (0 < r_k) \tag{7}$$

where $h|r_k$ which implies $h \leq r_k$.

Reversing the argument, we see from (7) that $r_k|r_{k-1}$ which implies $r_k|r_{k-2}$ by (6). Likewise

$$r_k|r_{k-2} \Rightarrow r_k|r_{k-3} \Rightarrow r_k|r_{k-4} \Rightarrow \cdots \Rightarrow r_k|r_2 \Rightarrow r_k|r_1 \Rightarrow r_k|n \Rightarrow r_k|m$$

the last two steps following from (5) and (4). Thus r_k is a common factor of both m and n, but because h is the *highest* common factor it must satisfy $h \geq r_k$. We have now proved that h satisfies both $h \leq r_k$ and $h \geq r_k$ from which we conclude $h = r_k$, the final remainder in Euclid's algorithm.

As a numerical example of its application, let us calculate the HCF of 2472 and 9216:

$$9216 = 3 \times 2472 + 1800$$

$$2472 = 1 \times 1800 + 672$$

$$1800 = 2 \times 672 + 456$$

$$672 = 1 \times 456 + 216$$

$$456 = 2 \times 216 + 24$$

$$216 = 9 \times 24$$

Thus 24, the last remainder, is the HCF of 2472 and 9216. It follows that the LCM of the two numbers is $2472 \times 9216/24 = 949{,}248$.

The HCF of the three numbers 2472, 9216 and 4616 is the HCF of 24 and 4616 which is easily shown by Euclid's algorithm to be 8. Since the HCFs of 2472 and 4616, and of 9216 and 4616, are both found to be 8 as well, the formula for the LCM of three numbers yields

$$\text{lcm}(2472{,}9216{,}4616) = \frac{2472 \times 9216 \times 4616 \times 8}{24 \times 8 \times 8} = 547716096$$

Finally, we derive from the Euclid algorithm a property of the HCF that is not particularly obvious from its definition, namely that there exist integers a and b (one of them being negative) such that $h = am + bn$. For, from (6) we have

$$h = r_k = r_{k-2} - c_k r_{k-1} = r_{k-2} - c_k(r_{k-3} - c_{k-1}r_{k-2}) = ur_{k-2} - c_k r_{k-3} = ur_{k-4} - vr_{k-3}$$
$$= wr_{k-4} - vr_{k-5} = \cdots$$

where $u = 1 + c_k c_{k-1}$, $v = c_k + uc_{k-2}$, $w = u(1 + c_{k-2}c_{k-3}) + c_k c_{k-3}$. Note that $c_k, u, v, w,$ etc. are all positive numbers so that one term in each step is positive and the other negative. The procedure continues as we work backwards through the algorithm to the final two equations, which according to (4) and (5), will take the form $h = \cdots = cn + ar_1 = am + bn$ with a, b and c representing integers to be determined.

Using the same numerical example ($m = 9216,\ n = 2472$) to illustrate the theory, we find that successive stages of the calculation give

$$m = 3n + 1800 \Rightarrow m - 3n = 1800$$

$$n = 1 \times (m - 3n) + 672 \Rightarrow 4n - m = 672$$

$$m - 3n = 2 \times (4n - m) + 456 \Rightarrow 3m - 11n = 456$$

$$4n - m = 1 \times (3m - 11n) + 216 \Rightarrow 15n - 4m = 216$$

$$3m - 11n = 2 \times (15n - 4m) + h \Rightarrow h = 11m - 41n.$$

Thus $a = 11$ and $b = -41$ in this example. Checking we see that

$$11 \times 9216 - 41 \times 2472 = 101{,}376 - 101{,}352 = 24$$

which verifies the stated property of the HCF.

Analytical form of the fundamental theorem

Now consider the expression

$$f_k(s) = \prod_{i=1}^{k} \frac{1}{1 - p_i^{-s}} \qquad (s > 1).$$

Since $(1 - x)^{-1} = \Sigma_{r=0}^{\infty} x^r$ for $|x| < 1$, and since $p_i \geq 2$, the expression under the product sign can be expanded in a convergent series, i.e.

$$f_k(s) = \prod_{i=1}^{k} \sum_{r=0}^{\infty} (p_i^{-s})^r \qquad (s > 1).$$

Suppose $k = 2$ for simplicity, then

$$f_2(s) = \sum_{r=0}^{\infty} (p_1^{-s})^r \sum_{t=0}^{\infty} (p_2^{-s})^t = \sum_{r=0}^{\infty} \sum_{t=0}^{\infty} (p_1^r p_2^t)^{-s}$$

where $p_1 = 2$ and $p_2 = 3$ of course. Clearly all numbers that have prime factors 2 and 3 raised to all possible combinations of powers will be included within the brackets of this double summation. For example,

$$1^{-s} = (2^0 \times 3^0)^{-s},\ 2^{-s} = (2^1 \times 3^0)^{-s},\ 3^{-s} = (2^0 \times 3^1)^{-s},\ 72^{-s} = (2^3 \times 3^2)^{-s},$$
$$(124{,}416)^{-s} = (2^9 \times 3^5)^{-s}$$

are five such terms in the sum defining $f_2(s)$. Thus we may write

$$f_2(s) = \sum_{p_1, p_2} n^{-s}$$

where the notation implies that summation is over all numbers n whose prime factors comprise every possible combination of powers of p_1 and p_2. By the Fundamental Theorem each n is uniquely expressed and can therefore only appear once in the summation. Likewise, $f_3(s)$ will be

the sum of all numbers with prime factors 2, 3 and 5 raised to all possible combinations of powers. In general, we have

$$f_k(s) = \sum_{p_1, p_2 \ldots p_k} n^{-s} > \sum_{n=1}^{p_k} n^{-s}$$

the inequality resulting from the fact that the numbers from 1 to p_k are already included in the first summation, as indicated by the first three terms in the example above for $f_2(s)$.

Since $s > 1$, the infinite series $\Sigma_{n=1}^{\infty} n^{-s}$ is convergent. It is in fact the well-known Riemann zeta function $\zeta(s)$. It follows from the inequality above that

$$f_k(s) > \sum_{n=1}^{p_k} n^{-s} = \sum_{n=1}^{\infty} n^{-s} - \sum_{n=p_k+1}^{\infty} n^{-s} > 0$$

which, on rearrangement, becomes

$$0 < \zeta(s) - f_k(s) < \sum_{n=p_k+1}^{\infty} n^{-s}. \qquad (8)$$

Now let $k \to \infty$ which means $p_k \to \infty$ as well. Then the right-hand side of (8) tends to 0 so that $f_\infty(s) = \lim_{k\to\infty} f_k(s) = \zeta(s)$, or with reference to the original definition of $f_k(s)$,

$$\zeta(s) = \prod_{i=1}^{\infty} \frac{1}{1 - p_i^{-s}}.$$

Hardy and Wright (*An Introduction to the Theory of Numbers*, 4th Edition, Oxford University Press, 1960) call this an analytical expression of the Fundamental Theorem of Arithmetic. It is an important result in the theory of primes as it relates them to the zeta function which has been extensively analysed.

Legendre's formula

It follows from the fundamental theorem that for any positive integer n

$$n! \equiv \prod_{m=1}^{n} m = \prod_{i=1}^{\infty} p_i^{k_i}$$

the expression on the right-hand side being the unique product of primes defining $n!$. Legendre's formula provides a way of determining the exponents $k_i \geq 0$ without the need to calculate $n!$ itself. To illustrate the method, we first consider a specific example with $n = 10, i = 2$ and look for the exponent k_2 of $p_2 = 3$. Clearly, every third number in the expression

$$10! = 1 \times 2 \times 3 \times 4 \times 5 \times 6 \times 7 \times 8 \times 9 \times 10$$

will have a prime factor 3. There are $\lfloor 10/3 \rfloor = 3$ such numbers, where the notation $\lfloor x \rfloor$, known as the floor function, means the greatest integer less than or equal to x or the 'integral part' of x, e.g. $\lfloor 2.4 \rfloor = 2,\ \lfloor 0.9 \rfloor = 0,\ \lfloor 1.0 \rfloor = 1$.

Note also that every ninth number is actually divisible by 3^2 thereby contributing an additional prime factor 3 each time. There are $\lfloor 10/3^2 \rfloor = 1$ such numbers in the product defining 10!. Likewise, every twenty-seventh number would be divisible by 3^3, but $\lfloor 10/3^3 \rfloor = 0$ indicating this is beyond the range of numbers in the product 10!. Clearly there will be no further contributions to k_2 from exponents greater than 3 and we deduce that

$$k_2 = \lfloor 10/3 \rfloor + \lfloor 10/3^2 \rfloor + \lfloor 10/3^3 \rfloor = 3 + 1 + 0 = 4$$

Applying the same arguments to $p_1 = 2, p_3 = 5$ and $p_4 = 7$ and stopping each time once we reach 0, we obtain

$$k_1 = \lfloor 10/2 \rfloor + \lfloor 10/2^2 \rfloor + \lfloor 10/2^3 \rfloor + \lfloor 10/2^4 \rfloor = 5 + 2 + 1 + 0 = 8$$

$$k_3 = \lfloor 10/5 \rfloor + \lfloor 10/5^2 \rfloor = 2 + 0 = 2$$

$$k_4 = \lfloor 10/7 \rfloor + \lfloor 10/7^2 \rfloor = 1 + 0 = 1$$

The next prime $p_5 = 11$ and all subsequent primes cannot be factors of 10! because they are all larger than the numbers in the product itself. Collecting the results obtained above, we see that the unique representation of 10! as a product of powers of primes is

$$10! = 3{,}628{,}800 = 2^8 \times 3^4 \times 5^2 \times 7$$

The generalisation of Legendre's formula is now obvious. Using the same arguments as in the specific example above, we have for any positive integer n

$$n! = \prod_{i=1}^{\infty} p_i^{k_i} \quad \text{where} \quad k_i = \sum_{r=1}^{\infty} \left\lfloor \frac{n}{p_i^r} \right\rfloor$$

Although the summation is infinite, we know that it will actually terminate after a finite number of terms when $p_i^r > n$, i.e. $\lfloor n/p_i^r \rfloor = 0$. At this point the infinite product also terminates because from then onwards it contributes only factors 1 to the product.

The product of n consecutive positive integers is divisible by n!

The proof of this simple result provides a nice application of Legendre's formula. We note first that since $x + y = (\lfloor x \rfloor + a) + (\lfloor y \rfloor + b)$, where $0 \le a < 1$ and $0 \le b < 1$, it follows that $\lfloor x + y \rfloor = \lfloor x \rfloor + \lfloor y \rfloor$ if $0 \le a + b < 1$ but $\lfloor x + y \rfloor = \lfloor x \rfloor + \lfloor y \rfloor + 1$ if $1 \le a + b < 2$. Hence we may assert in general that $\lfloor x \rfloor + \lfloor y \rfloor \le \lfloor x + y \rfloor$ for positive integers x and y.

Let the n successive integers begin with $m + 1$ $(m \ge 0)$. Thus we want to prove that $n! \,|(m + 1)(m + 2) \ldots (m + n)$, that is

$$\frac{(m+1)(m+2)\ldots(m+n)}{n!} = \frac{(m+n)!}{m!\,n!}$$

is a positive integer. We shall focus on one specific prime factor, p say, by writing $n! = p^k N$ where N represents the product of all the other prime factors of $n!$, and likewise $m! = p^j M$ and $(m+n)! = p^i P$, so that the ratio above becomes $p^i P/(p^j M p^k N) = p^{i-(j+k)} P/(MN)$, where P clearly contains all the prime factors appearing in M and N because it is greater than both of them. From Legendre's formula we have

$$i = \sum_{r=1}^{\infty} \left\lfloor \frac{m+n}{p^r} \right\rfloor, \quad j+k = \sum_{r=1}^{\infty} \left\lfloor \frac{m}{p^r} \right\rfloor + \sum_{r=1}^{\infty} \left\lfloor \frac{n}{p^r} \right\rfloor = \sum_{r=1}^{\infty} \left(\left\lfloor \frac{m}{p^r} \right\rfloor + \left\lfloor \frac{n}{p^r} \right\rfloor \right)$$

and since $\lfloor x \rfloor + \lfloor y \rfloor \leq \lfloor x+y \rfloor$ as noted earlier, we conclude that

$$\lfloor m/p^r \rfloor + \lfloor n/p^r \rfloor \leq \lfloor (m+n)/p^r \rfloor \quad \Rightarrow \quad j+k \leq i$$

This shows that the exponent $i-(j+k)$ of p is non-negative, thereby indicating that the factors p^j and p^k in the denominator $m!\,n!$ of the ratio always divide into the corresponding prime factor p^i in the numerator $(m+n)!$. Similar arguments apply to every other prime factor in the remaining factors M and N and we conclude that $(m+n)!$ is divisible by $m!\,n!$, which proves the original proposition.

Congruences and modular arithmetic

If a, b and m are integers, then the statement a is congruent to b modulo m means $m|(a-b)$ which is written formally as $a \equiv b \pmod{m}$. Another way of interpreting this definition is to observe that a and b will have the same remainders when divided by m. For if $a = q_1 m + r_1$ and $b = q_2 m + r_2$, where q_1 and q_2 are the quotients and r_1 and r_2 $(0 \leq r_{1,2} < m)$ are the remainders, then

$$a - b = (q_1 - q_2)m + (r_1 - r_2).$$

It follows that $a-b$ is divisible by m if and only if $r_1 = r_2$.

Clearly the definition $m|(a-b)$ is the same as stating $a - b = km$ where k is an integer. In other words, the congruence $a \equiv b \pmod{m}$ is equivalent to the equation $a = km + b$. The number b is called a residue of a modulo m. It can be regarded as what is left over after some multiple of m is subtracted from a. If $0 \leq b < m$ it is the same as the remainder (also called the least positive residue) introduced above, but other numbers, namely those that differ from b by a multiple of m (i.e. are congruent to b modulo m) are also residues. Modular arithmetic doesn't distinguish between such numbers; they are all regarded as equivalent and are said to belong to the same congruence class. For example, $100 \equiv 1 \pmod{9}$ but 100 is also congruent modulo 9 to $10, 19, 28, 37, -8, -17$ etc. all of which are members of the class containing 1 as its least positive residue.

It is obvious that $\{0, 1, 2, \cdots, m-1\}$ is a complete set of incongruent residues modulo m. Incongruent because being less than m they cannot differ from each other by some multiple of

m, and complete because all other integers will differ from one of the numbers in the set by some multiple of m and will therefore be congruent to that number. Each residue in the set defines its own congruence class modulo m; all other integers will be congruent to just one residue in the set and will therefore belong to the same congruence class as that residue. Each congruence class can be defined by any one of its members, not just the one stated in the given set, because all the members are congruent modulo m to each other. For example, we could have taken $\{1, 2, 3, \cdots, m\}$ as the defining set, or even $\{m, m+1, m+2, \cdots, 2m-1\}$.

We unconsciously use modular arithmetic in everyday life, the most obvious example being mod 12 arithmetic to measure time on a clock. When consulting a railway timetable or the departure time of a flight we may see the time given as 17.30. Without thinking, we would do a quick mental calculation in mod 12 arithmetic by subtracting 12 to obtain the least residue, and recognise the time in question as 5.30 pm. Similarly, if a doctor on a Monday appointment wants to see how the treatment is going after 25 days, we may be more concerned with the day of the week that will fall on, rather than the actual date, in case it clashes with some regular weekly commitment. Automatically we would switch to mod 7 arithmetic, subtract 21 from 25, and find it will be on a Friday, 4 days after Monday. These two examples correspond to the congruences $17 \equiv 5 \pmod{12}$ and $25 \equiv 4 \pmod{7}$ respectively.

It is a trivial exercise to prove that congruences obey most of the rules of ordinary algebra. For example (it is understood all congruences are $\bmod m$), if $a \equiv r$ and $b \equiv s$, then $a + b \equiv r + s$ and $ab \equiv rs$ because $a - r$, $b - s$, and $ab - rs = a(b - s) + s(a - r)$ are all divisible by m; if $a \equiv b$, and $b \equiv c$ then $a \equiv c$ because both a and c have the same remainder as b when divided by m; and if $a \equiv b$ then $ka \equiv kb$ because $m|k(a - b)$. An important exception, however, is that the congruence $ka \equiv kb$ does not necessarily reduce to $a \equiv b$ because the common factor k may absorb part of the division by m. On cancelling out the common factor 3 on each side of $27 \equiv 15 \pmod{6}$, for example, we are left with the incongruent relation $9 \not\equiv 5 \pmod{6}$. If k is coprime with m, however, then clearly $ka \equiv kb$ does indeed imply $a \equiv b$.

Another result we shall use in the next section is that the HCF of the modulus m and a residue b is the same for all numbers in the congruence class defined by b. Suppose $\text{hcf}(m, b) = d$ and let $b + km$ be some other number in the same congruence class as b, with $\text{hcf}(m, b + km) = d_1$. Since d is the HCF of m and b, $d|m$ and $d|b$ so that $d|(b + km)$. Thus $d \leq d_1$ since d_1 is the HCF of m and $b + km$. Likewise, $d_1|(b + km)$ and $d_1|m$ whence $d_1|(b + km - km)$, that is $d_1|b$ as well as $d_1|m$. But d is the HCF of m and b so that $d_1 \leq d$. The two conflicting inequalities are only satisfied if $d_1 = d$, that is any number in the congruence class defined by b has the same HCF with the modulus m as b itself. In particular, if $\text{hcf}(m, b) = 1$, in which case m and b are said to be coprime (or relatively prime), then all numbers in the same congruence class as b are coprime with m.

A number is divisible by 9 if and only if the sum of its digits is divisible by 9

This familiar arithmetical trick is known to many school students who otherwise have little interest in mathematics. Its proof, however, serves as an example of modular arithmetic and congruences.

We prove first by induction that $10^n \equiv 1 \pmod 9$ for all positive integers n. Assuming the statement is true, we note that $10^{n+1} - 1 = 10(10^n - 1) + 9$ is divisible by 9, or in other words $10^{n+1} \equiv 1 \pmod 9$. Thus $10^n \equiv 1 \pmod 9$ for all positive n since it is trivially true for $n = 1$. Hence $a10^n \equiv a \pmod 9$ and $a10^n + b10^m \equiv a + b \pmod 9$.

Now let a number c be written in the usual way as $a_n a_{n-1} \ldots a_1 a_0$, e.g. if $c = 9704$ we have $a_0 = 4,\ a_1 = 0,\ a_2 = 7, a_3 = 9$. This standard notation is, of course, shorthand for the expression $c = a_0 + a_1 10 + a_2 10^2 + \cdots + a_n 10^n$. By virtue of the result in the preceding paragraph we deduce that $c - a_0 \equiv a_1 + a_2 + \cdots + a_n \pmod 9$ or $9|\{c - (a_0 + a_1 + \cdots a_n)\}$. Thus if c is divisible by 9 then the sum of its digits must also be divisible by 9 and conversely, if 9 divides the sum of the digits of c, then c itself is divisible by 9.

The totient

The totient $\varphi(n)$ of a positive integer n is defined as the number of integers k $(1 \le k \le n)$ that are coprime with n. For example, there are eight such numbers $1, 5, 7, 11, 13, 17, 19, 23$ coprime with 24, so that $\varphi(24) = 8$ (note that since $\text{hcf}(1, n) = 1$, the number 1 is always counted in the totient). Likewise, $\varphi(25) = 20$ because only multiples of 5 are excluded. Since 23 is a prime number, all positive integers from 1 to 22 are prime to 23, so that $\varphi(23) = 22$. In fact, we can state in general that $\varphi(p) = p - 1$, where p is a prime. If $n > 1$, an alternative definition of $\varphi(n)$ is the number of integers k $(0 \le k \le n - 1)$ that are coprime with n. This replaces $k = n$ with $k = 0$, neither of which are coprime with n since $\text{hcf}(n, n) = \text{hcf}(0, n) = n$, so that the totient is unaffected.

Consider now the totient of the product of two relatively prime numbers. To illustrate how this is found in general, we begin with a numerical example. Let 4 and 9 be the relatively prime pair so that we want to calculate $\varphi(36)$. It is more convenient to use the alternative definition of the totient by arranging the numbers from 0 to 35 in a 4×9 array, as follows:

0	1	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

0	1	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

Clearly the numbers in the first row of the array are a complete set of residues modulo 9. Likewise, the second, third and subsequent rows are also complete sets of residues with the numbers in each column belonging to the same congruence class. Since $1, 2, 4, 5, 7, 8$ in the first row are coprime with 9 we have $\varphi(9) = 6$ and it follows from the result shown in the last section that all numbers in the six columns (or congruence classes) coloured red in the first array are also coprime with 9. On the same array on the right, the numbers coprime with 4 are coloured green or yellow. Note there are two such numbers in each column. Those that belong to

a red column in the array on the left, and are therefore coprime with 9 as well, are the ones coloured green.

Now numbers in the array will be coprime with 36 if and only if they are coprime with both 4 and 9. This is because the prime factors of 36 are $2 \times 2 \times 3 \times 3$ are the combined prime factors of 4 and 9 respectively. Therefore, if a number contains none of those factors it will be coprime with 36, and conversely any number coprime with 36 will not contain any of those same factors and will therefore be coprime with 4 and 9. Hence we seek those numbers, depicted in green in the right-hand array namely $1, 5, 7, 11, 13, 17, 19, 23, 25, 29, 31, 35$, which gives the result $\varphi(36) = 12 = \varphi(4)\varphi(9)$.

In general, we cannot rely on inspection to identify the coprime numbers; a more systematic approach is required. Accordingly, we note that any number in the first array can be expressed in the form $4q + r$ where q $(0 \leq q \leq 8)$ is the quotient when the number is divided by 4 and r $(0 \leq r \leq 3)$ is the remainder. Recalling that a coprime number in a red column, with number a $(0 \leq a \leq 8)$ in the first row, is given by $a + 9b$ $(0 \leq b \leq 3)$, we consider two such numbers $a + 9b_1$ and $a + 9b_2$ in the same red column. They can be expressed in the alternative forms

$$a + 9b_1 = 4q_1 + r_1 \qquad a + 9b_2 = 4q_2 + r_2$$

where, as we shall now show, $r_1 \neq r_2$. If $q_1 = q_2$ the remainders are necessarily different as otherwise there would only be one number defined. Assuming therefore that $q_1 \neq q_2$ and subtracting the two equations, we obtain

$$9(b_1 - b_2) = 4(q_1 - q_2) + r_1 - r_2$$

Now suppose $r_1 = r_2$ so that $9(b_1 - b_2) = 4(q_1 - q_2)$. Since 9 and 4 are coprime, they have no common factors, whence $9|(q_1 - q_2)$. But $0 < |q_1 - q_2| \leq 8$ so this is impossible. Thus we have a contradiction and conclude that $r_1 \neq r_2$, that is all the remainders in the column are different, ranging from 0 to 3 in some order. Let us take the red column with $a = 2$ as an example. Then $2 + 9b = 4q + r$ and as we let b run through its successive values from 0 to 3 we find that in the 1st, 2nd, 3rd and 4th rows respectively, $q = 0,\ r = 2;\ q = 2,\ r = 3;\ q = 5,\ r = 0;\ q = 7,\ r = 1$. The reminders are different in each of the four rows.

As proved in the previous discussion of congruence classes, $\text{hcf}(4, 4 + r) = \text{hcf}(4, r)$ and since $\text{hcf}(4, r) = 1$ when $r = 1$ or 3, there are just two numbers coprime with 4 in each column. In the third column with $a = 2$, the coprime numbers are in the 2nd and 4th rows and these are the numbers contributing to $\varphi(4)$. There are $\varphi(9)$ columns of numbers coprime with 9, each with $\varphi(4)$ numbers coprime with 4. Hence there are $\varphi(4)\varphi(9)$ numbers coprime with both 9 and 4, i.e. $\varphi(36) = \varphi(4)\varphi(9)$.

In the general case, let $\text{hcf}(m, n) = 1$ and consider the $m \times n$ array depicted below. As explained in the numerical example above, numbers are coprime with mn if and only if they are coprime with both m and n. The procedure for identifying these numbers follows that described in the numerical example and need only be given in outline here.

Representative columns and rows for general discussion are defined by the numbers k in the first row and jn in the first column respectively, where $0 \le j \le m-1$ and $0 \le k \le n-1$.

0	1	$\cdots$	k	$\cdots$	$n-2$	$n-1$
n	$1+n$	$\cdots$	$k+n$	$\cdots$	$2n-2$	$2n-1$
$2n$	$1+2n$	$\cdots$	$k+2n$	$\cdots$	$3n-2$	$3n-1$
$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$
jn	$1+jn$	$\cdots$	$k+jn$	$\cdots$	$jn-2$	$jn-1$
$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$
$(m-2)n$	$1+(m-2)n$	$\cdots$	$k+(m-2)n$	$\cdots$	$(m-1)n-2$	$(m-1)n-1$
$(m-1)n$	$1+(m-1)n$	$\cdots$	$k+(m-1)n$	$\cdots$	$mn-2$	$mn-1$

If $\text{hcf}(k,n)=1$ is coprime with n (corresponding to a red column in the previous example) then we know all the numbers in the column headed by k, which we shall call column k for simplicity, are also coprime with n. Conversely, if $\text{hcf}(k,n) \neq 1$, then none of the numbers in column k are prime to n. There are $\varphi(n)$ such columns. We also know from the arguments given in the numerical example that the number in column k that is located in the row identified by jn in the first column, can be expressed as $k+jn = q_j m + r_j$ where $\text{hcf}(k+jn,n) = \text{hcf}(r_j,m)$, with r_j taking on different values, ranging from 0 to $m-1$ in some order, for each value of j. Of these remainders, $\varphi(m)$ will be coprime with m.

Therefore, there are $\varphi(n)$ columns in which all the numbers are prime to n and in each of those columns $\varphi(m)$ of the numbers are also prime to m. Hence there are $\varphi(m)\varphi(n)$ numbers coprime with both m and n, i.e.

$$\text{hcf}(m,n)=1 \;\Leftrightarrow\; \varphi(mn)=\varphi(m)\varphi(n). \tag{9}$$

Note that if $m=p, n=q$, where p and q are both primes, then $\varphi(pq)=(p-1)(q-1)$.

Euler's theorem

Let $\text{hcf}(a,m)=1$ for $m \ge 1$ and let $n=\varphi(m)$. Among the complete set of residues modulo m, n of them, $\{a_1, a_2, \cdots, a_n\}$ say, will form a reduced set of residues that are coprime with m. For example, a complete set of residues modulo 10 is $\{0,1,2,3,4,5,6,7,8,9\}$ while the reduced set is $\{1,3,7,9\}$, the 4 numbers coprime with 10. The set $\{aa_1, aa_2, \cdots, aa_n\}$ is also a reduced set of residues because they are obviously all coprime with m and any two of them are incongruent modulo m. For if $aa_i \equiv aa_j$ $(i \neq j)$, then $a_i \equiv a_j$ which, since the residues a_i are incongruent in pairs, is only possible if $i=j$. (Note that although cancellation is not generally permissible in congruences, the a can be cancelled here because it is coprime with m.) Each residue in the set $\{aa_1, aa_2, \cdots, aa_n\}$ belongs to just one of the congruence classes defined by the original residues $\{a_1, a_2, \cdots, a_n\}$ but not necessarily in the same order. This is illustrated by choosing $a=13$ in the numerical example above with the reduced set of residues $\{1,3,7,9\}$. The new reduced set is $\{13,39,91,117\}$ where $13 \equiv 3, 39 \equiv 9, 91 \equiv 1, 117 \equiv 7 \pmod{10}$.

Since congruences can be multiplied like ordinary equations, the congruence

$$aa_1aa_2aa_3 \cdots \ aa_n \equiv a_1a_2a_3 \ \cdots \ a_n \pmod{m}$$

is valid regardless of which congruence classes the individual terms on the left-hand side belong to. It can be rewritten as

$$a^n a_1a_2a_3 \ \cdots \ a_n \equiv a_1a_2a_3 \ \cdots \ a_n \pmod{m}$$

The product $a_1a_2a_3 \ \cdots \ a_n$ and m are coprime because the residues in the reduced set are all coprime with m by definition and therefore their product is too. Thus the common factor $a_1a_2a_3 \ \cdots \ a_n$ can be cancelled in the congruence above, yielding $a^n \equiv 1 \pmod{m}$. This is Euler's theorem which can be stated more formally as

$$\text{hcf}(a, m) = 1 \quad \Rightarrow \quad a^{\varphi(m)} \equiv 1 \pmod{m}.$$

If $m = p$, a prime, then $\varphi(m) = p - 1$ and the condition $\text{hcf}(a, p) = 1$ is automatically satisfied. In this special case Euler's theorem becomes

$$a^p \equiv a \pmod{p}.$$

This result had already been discovered independently by Fermat and is often called Fermat's little theorem.

Invoking the multiplicative property of congruences again we may multiply the congruence in Euler's theorem by itself k times to obtain $\left(a^{\varphi(m)}\right)^k \equiv 1^k$, or

$$a^{k\varphi(m)} \equiv 1 \pmod{m}. \tag{10}$$

Encryption

We conclude by discussing an application of number theory to the encryption of data transmitted over the internet. It is an example of how even the purest branches of mathematics, which appear to be devoid of any practical use in the real world, can lead to unforeseen and surprising applications.

The RSA method, as it is called, relies on the fact that while it is easy to multiply two prime numbers together, the reverse procedure of factoring a composite number is much more difficult. Multiplying 83 by 59 is a trivial exercise in mental arithmetic but finding the two prime factors of 4897 is much more challenging. When very large prime numbers are involved, the factorisation of their product becomes near impossible.

We envisage a sender A who wants to transmit confidential data to a receiver B, even if a hacker E is able to monitor the two-way electronic traffic between them. It is conventional to give A, B and E the personal names Alice, Bob and Eve (for eavesdropper) respectively. In order to facilitate the encryption of data Alice wants to transfer to him, Bob first calculates $n = pq$ by multiplying two very large prime numbers p and q which are only known to him and then calculates

$$\varphi(n) = \varphi(pq) = \varphi(p)\varphi(q) = (p - 1)(q - 1)$$

using the multiplicative property (9) of the totient for coprime numbers. Finally, he chooses any number e coprime with $\varphi(n)$, and calculates d by solving the congruence $ed \equiv 1 \pmod{\varphi(n)}$, that is by finding d and k such that

$$ed = 1 + k\varphi(n).$$

The coprime condition is necessary because if $\text{hcf}\big(e, \varphi(n)\big) = h$, then $h|e$ and $h|\varphi(n)$ which implies $h|1$ in the equation above, i.e. $h = 1$.

In preparation for receiving encrypted data, Bob has the numbers p, q, n, e, d and $\varphi(n)$ at his disposal. He regards n and e as the public key (recall n is sufficiently large that it cannot be factored into its constituent primes p and q which are confidential to Bob) and sends them to Alice with no worry if Eve or others intercept them.

Alice's message is represented by the number x which could be numerical data or text converted into numerical form. On receipt of n and e, she converts x into an encrypted number y by computing

$$y \equiv x^e \pmod{n}.$$

and sends y to Bob. Again, there is no possibility of Eve converting y back into x because, as we shall see later, decryption requires a knowledge of d, and therefore of p and q, which are known only to Bob.

On the assumption that $\text{hcf}(x, n) = 1$, which is virtually certain given that the only factors of n are the large primes p and q, the extension (10) of Euler's theorem can now be used along with some routine algebra to recover x from y, as follows

$$y^d \equiv (x^e)^d = x^{ed} = x^{1+k\varphi(n)} = x \cdot x^{k\varphi(n)} \equiv x \cdot 1 \pmod{n}.$$

Hence, with his confidential key d, Bob can decode Alice's message by solving the congruence

$$x \equiv y^d \pmod{n}.$$

Before illustrating the theory of RSA encryption with a numerical example, we shall first show how seemingly formidable calculations in modular arithmetic, which arise in even the simplest numerical examples, can easily be performed on an ordinary calculator. In modulo n arithmetic, any integer a belongs to the congruent class defined by the remainder (or least positive residue) when a is divided by n, designated here by a_0 $(0 \leq a_0 \leq n - 1)$. Since $a \equiv a_0 \pmod{n}$, we have $a = kn + a_0$ and division by n gives

$$\frac{a}{n} = k + \frac{a_0}{n}$$

It is clear from the right-hand side of this equation that k is the integral part of a/n, and a_0/n is the decimal or fractional part because $0 \leq a_0 \leq n - 1$. It follows that

$$a_0 = n\left(\frac{a}{n} - \left\lfloor\frac{a}{n}\right\rfloor\right)$$

Hence to solve $x \equiv a \pmod{n}$ on a standard calculator, where a and n may be large numbers, we can reduce the congruence to $x \equiv a_0 \pmod{n}$ by entering the sequence of operations $(a \div n) = b;\ b - \lfloor b \rfloor = c;\ c \times n = a_0$. Many calculators have a separate key for the floor function which makes such calculations relatively easy. Some even have a modulo key that does the whole calculation in one step.

If the numbers are initially too large to be stored on the calculator, the congruence can be solved by breaking it down into component parts and solving each part separately. If $x \equiv ab \pmod{n}$ and if $x_1 \equiv a \pmod{n}$ and $x_2 \equiv b \pmod{n}$, then $x_1 x_2 \equiv ab \pmod{n}$, i.e. $x \equiv x_1 x_2 \pmod{n}$ and the two simpler congruences involving a and b can be solved first. Note that if $x \equiv a^{u+v}$, then we can choose $x_1 = a^u$ and $x_2 = a^v$. If $x \equiv a^{uv} \pmod{n}$, let $a^u \equiv b \pmod{n}$ and multiply this congruence by itself v times so that $(a^u)^v \equiv b^v \pmod{n}$ so that $x \equiv b^v \pmod{n}$ with the numerically simpler congruence $b \equiv a^u \pmod{n}$ being solved first.

Suppose Bob takes the two primes mentioned earlier, $p = 83$ and $q = 59$ (although in actual practice they would be much larger comprising up to several hundred digits), so that $n = 4897$ and $\varphi(n) = 82 \times 58 = 4756$. Clearly $\text{hcf}(7, 4756) = 1$, so Bob could choose $e = 7$. He now solves $7d \equiv 1 \pmod{4756}$, or $7d = 1 + 4756k$ for some k, obtaining $d = 1359$ with $k = 2$. He sends the two numbers 4897 and 7 to Alice representing his public key for n and e.

Alice proposes to send Bob the three-digit code on the back of her credit card which for simplicity we take to be 012. With the understanding that a two-digit number implies a leading zero, she sets $x = 12$ and computes $y \equiv 12^7 \pmod{4897}$ using the values of n and e sent to her by Bob. The least positive residue in the congruence class defined by $12^7 = 35{,}831{,}808$ is found by the procedure to calculate the remainder when 12^7 is divided by 4897 described in the introductory paragraph above, that is $\{(12^7 \div 4897) - \lfloor 12^7 \div 4897 \rfloor\} \times 4897 = 459$. Thus the number $y = 459$ is the encrypted version of $x = 12$ that Alice sends to Bob.

When Bob receives the coded message 459 he has to calculate $x \equiv 459^{1359} \pmod{4897}$ to recover Alice's credit card number. The calculation is simplified by letting $x \equiv x_1 x_2$ where $x_1 = 459^9$ and $x_2 = (x_1)^{150}$. (Here and henceforth all congruences are understood to be modulo 4897.) The calculator gives $\{(459^9 \div 4897) - \lfloor 459^9 \div 4897 \rfloor\} \times 4897 = 2802$ for the remainder of x_1 when divided by 4897. Hence $x_1 \equiv 2802$ and $x_2 \equiv 2802^{150} \equiv 1510$, the latter congruence resulting from a similar calculation for the remainder of $(2802^{150} \div 4897)$. Substitution of x_1 and x_2 in the original congruence for x gives $x \equiv 2802 \times 1510 = 4{,}231{,}020$, which reduces in the usual way to $\{(4231020 \div 4897 - \lfloor 4231020 \div 4897 \rfloor\} \times 4897 = 12$. Alice's original value of $x = 12$ has been successfully retrieved!

Almost magically, it seems, Alice's message has emerged from the jungle of calculations involving large numbers which arise even in this simple example. Bob now interprets the number 12 as indicating that the CVC number on Alice's credit card is 012 in the knowledge that it has been transmitted to him securely.