Chapter 5. Factorization of Integers

5.1 Definition: For $a, b \in \mathbf{Z}$ we say that a **divides** b (or that a is a **factor** of b, or that b is a **multiple of** a**), and we write** $a\big|b$**, when** $b = ak$ **for some** $k \in \mathbf{Z}$**.**

5.2 Theorem: *(Basic Properties of Divisors) Let* $a, b, c \in \mathbf{Z}$*. Then*

(1) $a|0$ *for all* $a \in \mathbf{Z}$ *and* $0|a \Longleftrightarrow a = 0$*,*
(2) $a|1 \Longleftrightarrow a = \pm 1$ *and* $1|a$ *for all* $a \in \mathbf{Z}$*.*
(3) If $a|b$ *and* $b|c$ *then* $a|c$*.*
(4) If $a|b$ *and* $b|a$ *then* $b = \pm a$*.*
(5) If $a|b$ *then* $|a| \leq |b|$*.*
(6) If $a|b$ *and* $a|c$ *then* $a\big|(bx + cy)$ *for all* $x, y \in \mathbf{Z}$*.*

Proof: Some of these properties hold in all rings while other properties are specific to $\mathbf{Z}$. Property (1) holds in all rings because when R is a ring and $a \in R$ we have $a \cdot 0 = 0$. Part, but not all, of Property (2) also holds in all rings. In any ring R we have $1|a$ for all $a \in R$ because $1 \cdot a = a$. Also in any ring R, because $1 \cdot 1 = 1$ and $(-1)(-1) = 1$ it follows that if $a = \pm 1$ then $a|1$. However, it is not the case that in every ring R and for all $a \in R$, if $a|1$ then $a = \pm 1$. For example, in the ring $\mathbf{Z}_8$ we have $1^2 = 3^2 = 5^2 = 7^2 = 1$ so that $1|1$, $3|1$, $5|1$ and $7|1$ (or equivalently $\pm 1|1$ and $\pm 3|1$). Let us prove that for all $a \in \mathbf{Z}$, if $a|1$ then $a = \pm 1$. Let $a \in \mathbf{Z}$ be arbitrary. Suppose that $a|1$. Choose $k \in \mathbf{Z}$ such that $1 = a \cdot k$. We claim that because $a \cdot k = 1$ it follows that either $a = k = 1$ or $a = k = -1$. Either $a < -1$ or $a = -1$ or $a = 0$ or $a = 1$ or $a > 1$. If $a > 1$ then

5.3 Theorem: *(The Division Algorithm) Let* $a, b \in \mathbf{Z}$ *with* $b \neq 0$*. Then there exist unique integers* q *and* r *such that*

$$a = qb + r \text{ and } 0 \leq r < |b|.$$

The integers q *and* r *are called the* **quotient** *and* **remainder** *when* a *is divided by* b*.*

Proof: We begin by proving that such integers q and r exist. Later we will show that the values of q and r are unique. Case 1: suppose that $b > 0$ and $a \geq 0$. Consider the sequence $0, b, 2b, 3b, \cdots$. Eventually the terms in the sequence become larger than a. Choose $q \geq 0$ so that $qb \leq a$ and $(q+1)b > a$. Let $r = a - qb$ so that $b = qa + r$. Since $qb \leq a$ we have $r = a - qb \geq 0$. Since $(q+1)b > a$, we have $qb + b > a$, and so $r = a - qb < b = |b|$.

Case 2: suppose that $b > 0$ and $a < 0$. Consider the sequence $0, -b, -2b, -3b, \cdots$. Eventually the terms in the sequence become smaller than a. Choose $p \geq 0$ so that $-(p-1)b > a$ and $-pb \leq a$. Let $q = -p$ so that $(q+1)b > a$ and $qb \leq a$. As above, we let $r = a - qb$ to get $a = qb + r$ and $0 \leq r < b = |b|$.

Case 3: suppose that $b < 0$ and $a \in \mathbf{Z}$. By the above two paragraphs, we can choose integers p and r so that $a = p|b| + r = -pb + r$ with $0 \leq r < |b|$, then we let $q = -p$ so that $a = qb + r$. In all three cases, we have shown that there exist integers q and r such that $a = qb + r$ and $0 \leq r < |b|$.

It remains to verify that the values of q and r are unique. Suppose that $a = qb + r$ with $0 \leq r < |b|$ and $a = pb + s$ with $0 \leq s < |b|$. Suppose, for a contradiction, that $r \neq s$ and say $r < s$ so that we have $0 \leq r < s < |b|$. Since $a = qb + r = pb + s$ we have $s - r = qb - pb = (q - p)b$ so that $b|(s - r)$. Since $b\big|(s - r)$ we have $|b| \leq |s - r| = s - r$ (by one of the basic properties of divisors). But since $s < |b|$ and $r \geq 0$ we have $s - r < |b|$ giving the desired contradiction. Thus we have $r = s$. Since $r = s$ and $s - r = (q - p)b$ we have $0 = (q - p)b$ hence $p = q$ (since $b \neq 0$).

5.4 Note: For $a, b \in \mathbf{Z}$, when we write $a = qb + r$ with $q, r \in \mathbf{Z}$ and $0 \leq r < |b|$, we have $b|a$ if and only if $r = 0$. Indeed if $r = 0$ then $a = qb$ so that $b|a$ and, conversely, if $b|a$ with say $a = pb = pb + 0$, then we must have $q = p$ and $r = 0$ by the uniqueness of the quotient and remainder.

5.5 Definition: Let $a, b \in \mathbf{Z}$. A **common divisor** of a and b is an integer d such that $d|a$ and $d|b$. When a and b are not both 0, we denote the **greatest common divisor** of a and b by $\gcd(a, b)$. For convenience, we also define $\gcd(0, 0) = 0$.

5.6 Theorem: *(Basic Properties of the Greatest Common Divisor) Let* $a, b, q, r \in \mathbf{Z}$.
(1) $\gcd(a, b) = \gcd(b, a)$.
(2) $\gcd(a, b) = \gcd(|a|, |b|)$.
(3) If $a|b$ *then* $\gcd(a, b) = |a|$. *In particular,* $\gcd(a, 0) = |a|$.
(4) If $b = qa + r$ *then* $\gcd(a, b) = \gcd(a, r)$.

Proof: The proof is left as an exercise.

5.7 Theorem: *(The Euclidean Algorithm With Back-Substitution) Let* a *and* b *be integers and let* $d = \gcd(a, b)$. *Then there exist integers* s *and* t *such that* $as + bt = d$. *The proof provides explicit procedures for finding* d *and for finding* s *and* t.

Proof: We can find d using the following procedure, called the **Euclidean Algorithm**. If $b|a$ then we have $d = |b|$. Otherwise, let $r_{-1} = a$ and $r_0 = b$ and use the division algorithm repeatedly to obtain integers q_i and r_i such that

$$\begin{array}{cc}
r_{-1} = a = q_1 b + r_1 & 0 < r_1 < |a| \\
r_0 = b = q_2 r_1 + r_2 & 0 < r_2 < r_1 \\
r_1 = q_3 r_2 + r_3 & 0 < r_3 < r_2 \\
\vdots & \vdots \\
r_{k-2} = q_k r_{k-1} + r_k & 0 < r_k < r_{k-1} \\
\vdots & \vdots \\
r_{n-2} = q_n r_{n-1} + r_n & 0 < r_n < r_{n-1} \\
r_{n-1} = q_{n+1} r_n + r_{n+1} & r_{n+1} = 0\,.
\end{array}$$

Since $r_{n-1} = q_{n+1} r_n$ we have $r_n \big| r_{n-1}$ so $\gcd(r_{n-1}, r_n) = r_n$. Since $r_{k-2} = q_k r_{k-1} + r_k$ we have $\gcd(r_{k-2}, r_{k-1}) = \gcd(r_{k-1}, r_k)$ and so

$$d = \gcd(a, b) = \gcd(b, r_1) = \gcd(r_1, r_2) = \cdots = \gcd(r_{n-2}, r_{n-1}) = \gcd(r_{n-1}, r_n) = r_n\,.$$

Having found d using the Euclidean algorithm, as above, we can find s and t using the following procedure, which is known as **Back-Substitution**. If $b|a$ so that $d = |b|$, then we can take $s = 0$ and $t = \pm 1$ to get $as + bt = d$. Otherwise, we let

$$s_0 = 1 \ , \ s_1 = -q_n \ , \text{ and } s_{\ell+1} = s_{\ell-1} - q_{n-\ell} s_\ell \text{ for } 1 \le \ell \le n-1$$

and then we can take $s = s_{n-1}$ and $t = s_n$ to get $as + bt = d$, because, writing $k = n - \ell$,

$$\begin{aligned}
d = r_n &= r_{n-2} - q_n r_{n-1} = s_1 r_{n-1} + s_0 r_{n-2} \\
&\ \ \vdots \\
&= \cdots = s_\ell r_{n-\ell} + s_{\ell-1} r_{n-\ell-1} = s_{n-k} r_k + s_{n-k-1} r_{k-1} \\
&= s_{n-k}\big(r_{k-2} - q_k r_{k-1}\big) + s_{n-k-1} r_{k-1} = \big(s_{n-k-1} - q_k s_{n-k}\big) r_{k-1} + s_{n-k} r_{k-2} \\
&= (s_{\ell-1} - q_{n-\ell} s_\ell) r_{n-\ell-1} + s_\ell r_{n-\ell-2} = s_{\ell+1} r_{n-\ell-1} + s_\ell r_{n-\ell-2} \\
&\ \ \vdots \\
&= \cdots = s_n r_0 + s_{n-1} r_{-1} = s_n b + s_{n-1} a \, .
\end{aligned}$$

5.8 Example: Let $a = 5151$ and $b = 1632$. Find $d = \gcd(a, b)$ and then find integers s and t so that $as + bt = d$.

Solution: The Euclidean Algorithm gives

$$\begin{aligned}
5151 &= 3 \cdot 1632 + 255 \\
1632 &= 6 \cdot 255 + 102 \\
255 &= 2 \cdot 102 + 51 \\
102 &= 2 \cdot 51 + 0
\end{aligned}$$

so $d = 51$. Using the quotients $q_1 = 3$, $q_2 = 6$ and $q_3 = 2$, Back-Substitution gives

$$\begin{aligned}
s_0 &= 1 \\
s_1 &= -q_3 = -2 \\
s_2 &= s_0 - q_2 s_1 = 1 - 6(-2) = 13 \\
s_3 &= s_1 - q_1 s_2 = -2 - 3(13) = -41 \, ,
\end{aligned}$$

so we take $s = s_2 = 13$ and $t = s_3 = -41$. (It is a good idea to check that indeed we have $(1632)(-41) + (5151)(13) = 51$).

5.9 Example: Let $a = 754$ and $b = -3973$. Find $d = \gcd(a, b)$ then find integers s and t such that $as + bt = d$.

Solution: The Euclidean Algorithm gives

$3973 = 5{\cdot}754{+}203$, $754 = 3{\cdot}203{+}145$, $203 = 1{\cdot}145{+}58$, $145 = 2{\cdot}58{+}29$, $58 = 2{\cdot}29{+}0$

so that $d = 29$. Then Back-Substitution gives rise to the sequence

$$1 \ , \ \ -2 \ , \ \ 3 \ , \ \ -11 \ , \ \ 58$$

so we have $(754)(58) + (3973)(-11) = 29$, that is $(754)(58) + (-3973)(11) = 29$. Thus we can take $s = 58$ and $t = 11$.

5.10 Theorem: *(More Properties of the Greatest Common Divisor) Let $a, b, c, d \in \mathbf{Z}$.*
(1) If $c|a$ and $c|b$ then $c \Big| \gcd(a, b)$.
(3) We have $\gcd(a, b) = 1$ if and only if there exist $x, y \in \mathbf{Z}$ such that $ax + by = 1$.
(4) If $d = \gcd(a, b) \neq 0$ then $\gcd\left(\frac{a}{d}, \frac{b}{d}\right) = 1$.
(5) If $a|bc$ and $\gcd(a, b) = 1$ then $a|c$.

Proof: We prove Part (5). Suppose that $a|bc$ and $\gcd(a, b) = 1$. Since $a|bc$ we can choose $k \in \mathbf{Z}$ so that $bc = ak$. Since $\gcd(a, b) = 1$, by the Euclidean Algorithm with Back-Substitution, we can choose $s, t \in \mathbf{Z}$ with $as + bt = 1$. Then we have

$$c = c \cdot 1 = c(as + bt) = acs + bct = acs + akt = a(cs + kt),$$

and so $a|c$, as required.

5.11 Definition: A **diophantine equation** is a polynomial equation in which the variables represent integers. Some diophantine equations are fairly easy to solve while others can be extremely difficult.

5.12 Theorem: *(Linear Diophantine Equations) Let $a, b, c \in \mathbf{Z}$ with $(a, b) \neq (0, 0)$. Let $d = \gcd(a, b)$ and note that $d \neq 0$. Consider the Diophantine equation $ax + by = c$.*
(1) The equation has a solution $(x, y) \in \mathbf{Z}^2$ if and only if $d|c$, and
(2) if $(u, v) \in \mathbf{Z}^2$ is one solution to the equation then the general solution is given by

$$(x, y) = (u, v) + k\big(-\tfrac{b}{d}, \tfrac{a}{d}\big) \text{ for some } k \in \mathbf{Z}.$$

Proof: Suppose that the equation $ax+by = c$ has a solution $(x, y) \in \mathbf{Z}^2$. Choose $(s, t) \in \mathbf{Z}^2$ so that $as + bt = c$. Since $d|a$ and $d|b$, it follows that $d|(ax + by)$ for all $x, y \in \mathbf{Z}$, so in particular $d|(as+bt)$, that is $d|c$. Conversely, suppose that $d|c$, say $c = d\ell$ with $\ell \in \mathbf{Z}$. Use the Euclidean Algorithm with Back-Substitution to find $s, t \in \mathbf{Z}$ such that $as + bt = d$. Multiply by ℓ to get $a(s\ell) + b(t\ell) = d\ell = c$. Thus we can take $x = e\ell$ and $y = t\ell$ to obtain a solution $(x, y) \in \mathbf{Z}^2$ to the equation $ax + by = c$. This proves Part (1)

Now suppose that $(u, v) \in \mathbf{Z}^2$ is a solution to the given equation, so we have $au+bv = c$. To prove Part (2), we need to prove that for all $k \in \mathbf{Z}$, if we let $(x, y) = (u, v) + k\big(-\frac{b}{d}, \frac{a}{d}\big)$ then (x, y) is a solution to $ax + by = c$ and, conversely, that if (x, y) is a solution then there exists $k \in \mathbf{Z}$ such that $(x, y) = (u, v) + k\big(-\frac{b}{d}, \frac{a}{d}\big)$.

Let $k \in \mathbf{Z}$ and let $(x, y) = (u, v) + k\big(-\frac{b}{d}, \frac{a}{d}\big)$. Then $x = u - \frac{kb}{d}$ and $y = v + \frac{ka}{d}$ and so

$$ax + by = a\big(u - \tfrac{kb}{d}\big) + b\big(v + \tfrac{ka}{d}\big) = (au + bv) - \tfrac{kab}{d} + \tfrac{kab}{d} = au + bv = c.$$

Conversely, let (x, y) be a solution to the given equation, so we have $ax + by = c$. Suppose that $a \neq 0$ (we leave the case $a = 0$ as an exercise). Since $ax + by = c$ and $au + bu = c$ we have $ax + by = au + bv$ and so $a(x - u) = -b(y - v)$. Divide both sides by d to get $\frac{a}{d}(x - u) = -\frac{b}{d}(y - v)$. Since $\frac{a}{d} \Big| \frac{b}{d}(y - v)$ and $\gcd\left(\frac{a}{d}, \frac{b}{d}\right) = 1$, it follows that $\frac{a}{d} \Big| (y - v)$. Choose $k \in \mathbf{Z}$ so that $y - v = \frac{ka}{d}$. Since $a \neq 0$ and $a(x - u) = -b(y - v) = -\frac{kab}{d}$, we have $x - u = -\frac{kb}{d}$ and so $(x, y) = (u, v) + k\big(-\frac{b}{d}, \frac{a}{d}\big)$, as required.

5.13 Example: Let $a = 426$, $b = 132$ and $c = 42$. Find all $x, y \in \mathbf{Z}$ such that $ax + by = c$.

Solution: The Euclidean Algorithm gives

$$426 = 3 \cdot 132 + 30 \ , \ 132 = 4 \cdot 30 + 12 \ , \ 30 = 2 \cdot 12 + 6 \ , \ 12 = 2 \cdot 6 + 0$$

so that $d = \gcd(a, b) = 6$. Note that $d|c$, indeed $c = d\ell$ with $\ell = 7$, so a solution does exist. Back-Substitution gives the sequence

$$1 \ , \ -2 \ , \ 9 \ , \ -29$$

so we have $a(9) + b(-29) = d$. Multiply by $\ell = 7$ to get $a(63) + b(-203) = c$, so one solution is given by $(x, y) = (63, -203)$. Since $\frac{a}{d} = \frac{426}{6} = 71$ and $\frac{b}{d} = \frac{132}{6} = 22$, The general solution is $(x, y) = (63, -203) + k(-22, 71)$.

5.14 Exercise: Let $a = 4123$, $b = 17689$ and $c = 798$. Find all $x, y \in \mathbf{Z}$ with $0 \leq y \leq 100$ such that $ax + by = c$.

5.15 Example: A **Pythagorean triple** is a solution (x, y, z) with $x, y, z \in \mathbf{Z}^+$ to the equation $x^2 + y^2 = z^2$. Note that when (x, y, z) is a Pythagorean triple with $z \neq 0$, we have $\left(\frac{x}{z}\right)^2 + \left(\frac{y}{z}\right)^2 = 1$ so that the point $\left(\frac{x}{z}, \frac{y}{z}\right)$ is a point on the unit circle with rational coordinates. Let S be the unit circle $\{(x, y) | x^2 + y^2 = 1\}$ and let $T = S \setminus \{(0, 1)\}$. The **stereographic projection** from T to $\mathbf{R}$ is the function $F : T \to \mathbf{R}$ defined as follows: given $(x, y) \in T$, let L be the line through $(0, 1)$ and (x, y), and define $F(x, y) = u$ where $(u, 0)$ is the point of intersection of L with the x-axis. Let us find a formula for F and a formula for its inverse $G : \mathbf{R} \to T$.

Given $(x, y) \in T$, the line L from $(0, 1)$ to (x, y) is given parametrically by $(u, v) = (0, 1) + t((x, y) - (0, 1)) = (tx, 1 + t(y - 1))$. This line meets the x-axis when $0 = v = 1 + t(y - 1)$, that is when $t = \frac{1}{1-y}$, and the resulting point of intersection is at $(u, v) = (tx, 1 + t(y - 1)) = \left(\frac{x}{1-y}, 0\right)$. Thus the map F is given by $u = F(x, y) = \frac{x}{1-y}$.

Given a point $(u, 0)$ on the x-axis, the line M through $(0, 1)$ and $(u, 0)$ is given parametrically by $(x, y) = (0, 1) + t((u, 0) - (0, 1)) = (tu, 1 - t)$. The point $(x, y) = (tu, 1 - t)$ lies on S when $1 = x^2 + y^2 = (tu)^2 + (1 - t)^2 = t^2u^2 + 1 - 2t + t^2$, that is when $(u^2 + 1)t^2 = 2t$, or equivalently when $t = 0$ or $t = \frac{2}{u^2+1}$. When $t = 0$ the resulting point is $(x, y) = (tu, 1 - t) = (0, 1)$ and when $t = \frac{2}{u^2+1}$ the resulting point is $(x, y) = (tu, 1 - t) = \left(\frac{2u}{u^2+1}, \frac{u^2-1}{u^2+1}\right)$. Thus the inverse map G is given by $(x, y) = G(u) = \left(\frac{2u}{u^2+1}, \frac{u^2-1}{u^2+1}\right)$.

Notice that if $(x, y) \in T$ with $x, y \in \mathbf{Q}$ then $u = F(x, y) \in \mathbf{Q}$ and that, conversely, if $u \in \mathbf{Q}$ then $(x, y) = G(u) \in \mathbf{Q}^2$. It follows that we have a bijective correspondence between $T \cap \mathbf{Q}^2$ and $\mathbf{Q}$ given by $F : T \cap \mathbf{Q}^2 \to \mathbf{Q}$ and $G : \mathbf{Q} \to T \cap \mathbf{Q}^2$. Thus every element in $T \cap \mathbf{Q}^2$ is of the form

$$G\left(\tfrac{s}{t}\right) = \left(\tfrac{2(s/t)}{(s/t)^2+1}, \tfrac{(s/t)^2-1}{(s/t)^2+1}\right) = \left(\tfrac{2st}{s^2+t^2}, \tfrac{s^2-t^2}{s^2+t^2}\right)$$

for some $s, t \in \mathbf{Z}$ with $t \neq 0$ and $\gcd(s, t) = 1$.

After doing some additional work (which involves considering the case in which $s + t$ is even and the case in which $s + t$ is odd, and in the former case replacing s and t by $s' = \frac{s+t}{2}$ and $t' = \frac{s-t}{2}$) one can verify that every Pythagorean triple (x, y, z), after possibly interchanging x and y, is of the form $r\left(2st \, , \, s^2 - t^2 \, , \, s^2 + t^2\right)$ for some $r, s, t \in \mathbf{Z}$ with $\gcd(s, t) = 1$ and $s + t$ odd.

5.16 Definition: Let n be a positive integer. We say that n is a **prime number** when $n \geq 2$ and n has no factor $a \in \mathbf{Z}$ with $1 < a < n$. We say that n is **composite** when $n \geq 2$ and n is not prime, that is when n does have a factor $a \in \mathbf{Z}$ with $1 < a < n$.

5.17 Theorem: *(Basic Properties of Primes) Let p be a prime number.*

(1) For all $a \in \mathbf{Z}$ we have $\gcd(a, p) \in \{1, p\}$ with $\gcd(a, p) = p$ if and only if $p|a$.

(2) For all $a, b \in \mathbf{Z}$, if $p|ab$ then either $p|a$ or $p|b$.

Proof: The proof is left as an exercise. Part (2) follows from Part (5) of Theorem 5.10.

5.18 Theorem: *Every integer $n \geq 2$ has a prime factor. Every composite integer $n \geq 2$ has a prime factor p with $p \leq \sqrt{n}$.*

Proof: Let $n \geq 2$. Suppose, inductively, that every integer k with $2 \leq k < n$ has a prime factor. If n is prime, then n is a prime factor of itself, so n has a prime factor. Suppose that n is composite. Let a be a factor of n with $1 < a < n$. By the induction hypothesis, a has a prime factor. Let p be a prime factor of a. Since $p|a$ and $a|n$ we have $p|n$, and so p is a prime factor of n. It follows, by induction, that every integer $n \geq 2$ has a prime factor.

Now suppose that n is composite. Write $n = ab$ where $a, b \in \mathbf{Z}$ with $1 < a \leq b < n$. Note that $a \leq \sqrt{n}$ because if we had $a > \sqrt{n}$ then we would also have $b \geq a > \sqrt{n}$ so that $n = ab > \sqrt{n}\sqrt{n} = n$ which is impossible. Let p be a prime factor of a. Since $p|a$ and $a|n$ we have $p|n$ so that p is a prime factor of n. Since $p|a$ and $a \leq \sqrt{n}$ we have $p \leq a \leq \sqrt{n}$.

5.19 Note: Given an integer $n \geq 2$, we can list all primes p with $p \leq n$ using the following procedure, which is called the **Sieve of Eratosthenes**. We begin by listing all the integers from 1 to n, and we cross off the number 1 (1 is a unit; it is not a prime). We circle the smallest remaining number p_1 (namely $p_1 = 2$, which is prime) then we cross off all other multiples of p_1 (which are composite). We circle the smallest remaining number p_2 (namely $p_2 = 3$, which is prime) then we cross off all other multiples of p_2 (which are all composite). At the k^{th} step of the procedure, when we circle the smallest remaining number p_k, it must be prime because if p_k was composite then it would have a prime factor p_i with $p_i < p_k$, but we have already found all primes $p_i < p_k$ and we have already crossed off all their multiples. We continue the procedure until we have circled a prime p_ℓ with $p_\ell \geq \sqrt{n}$ and crossed off its multiples. At this stage we circle all of the remaining numbers in the list because they are all prime. Indeed, if a remaining number m was composite then it would have a prime factor p with $p \leq \sqrt{m} \leq \sqrt{n}$, but we have already found all primes p with $\leq \sqrt{n}$ and crossed off all their multiples.

5.20 Exercise: Use the Sieve of Eratosthenes to list all primes p with $p \leq 100$.

5.21 Theorem: *(Euclid) There exist infinitely many prime numbers.*

Proof: Suppose, for a contradiction, that there exist finitely many prime numbers. Let $p_1, p_2, \cdots, p_\ell$ be all of the prime numbers. Consider the number $n = p_1 p_2 \cdots p_\ell + 1$. By Theorem 5.11, the number n has a prime factor and so $p_k|n$ for some index k. But p_k is not a factor of n because when we write $n = qp_k + r$ as in the Division Algorithm, we find that the remainder is $r = 1 \neq 0$ (and the quotient is $q = \prod_{i \neq k} p_i$).

5.22 Example: Note that there exist arbitrarily large gaps between consecutive prime numbers because, given a positive integer $m \geq 2$, we have $2\big|(m!+2)$, $3\big|(m!+3)$, $4\big|(m!+4)$ and so on, so the consecutive numbers $m!+2$, $m!+3$, $m!+4$, $\cdots$, $m!+m$ are all composite.

5.23 Remark: Here are a few facts about prime numbers which are difficult to prove.

(1) Bertrand's Postulate: for every integer $n \geq 1$ there exists a prime p with $n \leq p \leq 2n$.

(2) Dirichlet's Theorem: for all positive integers a, b with $\gcd(a, b) = 1$, there exist infinitely many primes of the form $p = a + kb$ for some $k \in \mathbf{N}$.

(3) The Prime Number Theorem: for $x \in \mathbf{R}$, let $\pi(x)$ be the number of primes p with $p \leq x$. Then $\lim_{x\to\infty} \dfrac{\pi(x)}{x/\ln x} = 1$.

5.24 Remark: Here are a few statements about prime numbers which are conjectured to be true, but for which no proof has, as yet, been found.

(1) Legendre's Conjecture: for every $n \in \mathbf{Z}^+$ there exists a prime p with $n^2 \leq p \leq (n+1)^2$.

(2) Goldbach's Conjecture: every even integer $n \geq 4$ is the sum of two prime numbers.

(3) Twin Primes Conjecture: there exist infinitely many p for which p and $p+2$ are prime.

(4) The $n^2 + 1$ Conjecture: there exist infinitely many primes $p = n^2 + 1$ with $n \in \mathbf{Z}^+$.

(5) Mersenne Primes Conjecture: there exist infinitely many primes $p = 2^k - 1$ with $k \in \mathbf{Z}^+$.

(6) Fermat Primes Conjecture: there exist finitely many primes $p = 2^k + 1$ with $k \in \mathbf{N}$.

5.25 Theorem: *(The Unique Factorization Theorem) Every integer $n \geq 2$ can be written uniquely in the form $n = \prod_{k=1}^{\ell} p_k = p_1 p_2 \cdots p_\ell$ where $\ell \in \mathbf{Z}^+$ and the p_k are primes with $p_1 \leq p_2 \leq \cdots \leq p_\ell$.*

Proof: First we prove the existence of such a factorization. Let n be an integer with $n \geq 2$ and suppose, inductively, that every integer k with $2 \leq k < n$ can be written in the required form. If n is prime then we can write $n = \prod_{k=1}^{\ell} p_k = p_1$ with $\ell = 1$ and $p_1 = n$. Suppose that n is composite. Write $n = ab$ where $a, b \in \mathbf{Z}$ with $1 < a < n$ and $1 < b < n$. By the induction hypothesis, we can write $a = q_1 q_2 \cdots q_\ell$ and $b = r_1 r_2 \cdots r_m$ where $\ell, m \in \mathbf{Z}^+$ and the p_i and q_i are primes with $p_1 \leq p_2 \leq \cdots \leq p_\ell$ and $q_1 \leq q_2 \leq \cdots \leq q_m$. Then $n = q_1 q_2 \cdots q_\ell r_1 r_2 \cdots r_m = p_1 p_2 \cdots p_{\ell+m}$ where the ordered $(\ell+m)$-tuple $(p_1, p_2, \cdots, p_{\ell+m})$ is obtained from the ordered $(\ell+m)$-tuple $(q_1, q_2, \cdots, q_\ell, r_1, r_2, \cdots, r_m)$ by rearranging the terms into non-decreasing order.

Let us prove uniqueness. Suppose that $n = p_1 p_2 \cdots p_\ell = q_1 q_2 \cdots q_m$ where $\ell, m \in \mathbf{Z}^+$ and the p_i and q_j are primes with $p_1 \leq p_2 \leq \cdots \leq p_\ell$ and $q_1 \leq q_2 \leq \cdots \leq q_m$. We need to prove that $\ell = m$ and that $p_i = q_i$ for every index i. Since $n = p_1 p_2 \cdots p_\ell$ we see that $p_1 | n$ and so $p \big| q_1 q_2 \cdots q_m$. By applying Part (2) of Theorem 5.12 repeatedly, it follows that $p_1 \big| q_i$ for some index i. Since $p_1 | q_i$ and q_i is prime, we must have $p_1 \in \{\pm 1, \pm q_i\}$. Since p_1 is prime, we have $p_1 > 1$. Since $p_1 > 1$ and $p_1 \in \{\pm 1, \pm q_i\}$ it follows that $p_1 = q_i$. A similar argument shows that $q_1 = p_j$ for some index j. Since $p_1 = q_i \geq q_1 = p_j \geq p_1$, it follows that $p_1 = q_1$.

Since $p_1 p_2 \cdots p_\ell = q_1 q_2 \cdots q_m$ and $p_1 = q_1$, we can divide both sides by p_1 to get $p_2 p_3 \cdots p_\ell = q_2 q_3 \cdots q_m$. By repeating the above argument, we can show that $p_2 = q_2$, then we can divide both sides by $p_2 = q_2$ to get $p_3 \cdots p_\ell = q_3 \cdots q_m$ and so on.

If we had $\ell \neq m$, say $\ell < m$, repeating the above procedure would eventually yield $p_\ell = q_\ell q_{\ell+1} \cdots q_m$ with $p_\ell = q_\ell$ and then $1 = q_{\ell+1} \cdots q_m$ which is not possible since each $q_i > 1$. Thus we must have $\ell = m$ and repeating the above procedure gives $p_i = q_i$ for all indices i, as required.

5.26 Note: Here are two alternate ways of expressing the above theorem.

(1) Every integer $n \geq 2$ can be written uniquely in the form $n = \prod_{i=1}^{\ell} {p_i}^{m_i} = {p_1}^{m_1} \cdots {p_\ell}^{m_\ell}$ where $\ell \in \mathbf{Z}^+$ and the p_i are distinct primes with $p_1 < p_2 < \cdots < p_\ell$ and each $m_i \in \mathbf{Z}^+$.

(2) Given distinct primes $p_1, p_2, \cdots, p_\ell$, every $n \in \mathbf{Z}^+$ whose prime factors are included in $\{p_1, \cdots, p_\ell\}$ can be written uniquely in the form $n = \prod_{i=1}^{\ell} {p_i}^{m_i} = {p_1}^{m_1} \cdots {p_\ell}^{m_\ell}$ with $m_i \in \mathbf{N}$.

5.27 Theorem: *(Unique Factorization and Divisors) Let $n = {p_1}^{m_1}{p_2}^{m_2} \cdots {p_\ell}^{m_\ell}$ where $\ell \in \mathbf{Z}^+$, the p_i are distinct primes, and each $m_i \in \mathbf{N}$. Then the positive divisors of n are the numbers of the form $a = {p_1}^{j_1}{p_2}^{j_2} \cdots {p_\ell}^{j_l}$ where each $j_i \in \mathbf{Z}$ with $0 \leq j_i \leq m_i$.*

Proof: Suppose that $n = {p_1}^{m_1}{p_2}^{m_2} \cdots {p_\ell}^{m_\ell}$ and $a = {p_1}^{j_1}{p_2}^{j_2} \cdots {p_\ell}^{j_\ell}$ where $p_1, p_2, \cdots, p_\ell$ are distinct primes and $0 \leq j_i \leq m_i$ for all indices i. Let $b = {p_1}^{k_1}{p_2}^{k_2} \cdots {p_\ell}^{k_\ell}$ where $k_i = m_i - j_i$ (note that $k_i \geq 0$ since $j_i \leq m_i$). Then

$$ab = ({p_1}^{j_1} \cdots {p_\ell}^{j\ell})({p_1}^{k_1} \cdots {p_\ell}^{k_\ell}) = {p_1}^{j_1+k_1} \cdots {p_\ell}^{j_\ell+k_\ell} = {p_1}^{m_1} \cdots {p_\ell}^{m_\ell} = n$$

and so $a|n$.

Conversely, suppose that $n = {p_1}^{m_1}{p_2}^{m_2} \cdots {p_\ell}^{m_\ell}$, as above, and let a be a positive divisor of n. Let p be any prime factor of a. Since $p|a$ and $a|n$ we have $p|n$. Since $p|n$ and $n = {p_1}^{m_1}{p_2}^{m_2} \cdots {p_\ell}^{m_\ell}$ we have $p|p_i$ for some index i. Since p and p_i are both prime and $p|p_i$, we have $p = p_i$. This proves that every prime factor of a is among the primes $p_1, p_2, \cdots, p_\ell$. It follows that a can be written in the form $a = {p_1}^{j_1}{p_2}^{j_2} \cdots {p_\ell}^{j_\ell}$ with each $j_i \in \mathbf{N}$. It remains to show that $j_i \leq m_i$.

Since $a|n$ we can choose $b \in \mathbf{Z}$ so that $n = ab$. Since n and a are positive, so is b. Since b is a positive factor of n, the above argument shows that every prime factor of b is among the primes $p_1, p_2, \cdots, p_\ell$ and so we can write $b = {p_1}^{k_1}{p_2}^{k_2} \cdots {p_\ell}^{k_\ell}$ for some $k_i \in \mathbf{N}$. Since $n = ab$ we have

$${p_1}^{m_1}{p_2}^{m_2} \cdots {p_\ell}^{m_\ell} = n = ab = ({p_1}^{j_1} \cdots {p_\ell}^{j_\ell})({p_1}^{k_1} \cdots {p_\ell}^{k_\ell}) = {p_1}^{j_1+k_1} \cdots {p_\ell}^{j_\ell+k_\ell}.$$

By the uniqueness of prime factorization, it follows that $m_i = j_i + k_i$ for all indices i. Since $k_i \geq 0$ it follows that $j_i = m_i - k_i \leq m_i$, as required.

5.28 Definition: For $a, b \in \mathbf{Z}$, a **common multiple** of a and b is an integer m such that $a|m$ and $b|m$. When a and b are nonzero, we define $\text{lcm}(a, b)$ to be the smallest positive common multiple of a and b. For convenience, we also define $\text{lcm}(a, 0) = \text{lcm}(0, a) = 0$ for $a \in \mathbf{Z}$.

5.29 Theorem: *Let $a = \prod_{i=1}^{\ell} {p_i}^{j_i}$ and $b = \prod_{i=1}^{\ell} {p_i}^{k_i}$ where $\ell \in \mathbf{Z}^+$, the p_i are distinct primes, and $j_i, k_i \in \mathbf{N}$. Then*

(1) $\gcd(a, b) = \prod_{i=1}^{\ell} {p_i}^{\min\{j_i, k_i\}}$,

(2) $\text{lcm}(a, b) = \prod_{i=1}^{\ell} {p_i}^{\max\{j_i, k_i\}}$, *and*

(3) $\gcd(a, b) \cdot \text{lcm}(a, b) = ab$.

Proof: The proof is left as an exercise.

5.30 Definition: For a prime p and a positive integer n, the **exponent** of p in (the prime factorization of) n, denoted by $e(p,n)$, is defined as follows. We write n in the form $n = {p_1}^{m_1}{p_2}^{m_2}\cdots{p_\ell}^{m_\ell}$ where the p_i are distinct primes and each $m_i \in \mathbf{N}$, then we define $e(p,n) = m_i$ if $p = p_i$ and we define $e(p,n) = 0$ if $p \neq p_i$ for any index i.

5.31 Exercise: Show that $e(p,n!) = \lfloor \frac{n}{p} \rfloor + \lfloor \frac{n}{p^2} \rfloor + \lfloor \frac{n}{p^3} \rfloor + \cdots$ and that $\lfloor \frac{n}{p^{k+1}} \rfloor = \left\lfloor \lfloor \frac{n}{p^k} \rfloor \Big/ p \right\rfloor$.

5.32 Example: Since $e(5,100!) = \lfloor \frac{100}{5} \rfloor + \lfloor \frac{100}{25} \rfloor + \lfloor \frac{100}{125} \rfloor + \cdots = 20 + 4 + 0 = 24$ and $e(2,100!) > 24$, it follows that the number 100! ends with exactly 24 zeros in its decimal representation.

5.33 Definition: For a positive integer n, we write $\tau(n)$ to denote the number of positive divisors of n, we write $\sigma(n)$ to denote the sum of the positive divisors of n, and we write $\rho(n)$ to denote the product of the positive divisors of n.

5.34 Exercise: Let $n = \prod_{i=1}^{\ell} {p_i}^{k_i}$ where $p_1, p_2, \cdots, p_\ell$ are distinct primes and each $k_i \in \mathbf{N}$. Show that $\tau(n) = \prod_{i=1}^{\ell} (k_i + 1)$, $\sigma(n) = \prod_{i=1}^{\ell} \frac{{p_i}^{k_i - 1} - 1}{p_i - 1}$ and $\rho(n) = n^{\tau(n)/2}$.