YQL's Notes: Introduction to Abstract Algebra

† Contents and problems from http://fmwww.bc.edu/gross/MT216/MT216_notes.pdf

† This is my first systematic LaTeX math note. Please expect typos and other LaTeX-related mistakes. Once spotted, they will be fixed.

March 29, 2020

Catalogue

1 Mathematical Induction

1.1 The Principle of Induction

Theorem 1.1.1. Let a be a natural number, and let $P(n)$ be a statement about n for each integer $n \geqslant a$. If

(1) $P(a)$ is true, and

(2) for every integer $k \geqslant a$, $P(k) \implies P(k+1)$,

then $P(k)$ is true for every integer $k \geqslant a$.

Problem 1.1.1 (1.2.29). Show that for $n \in \mathbb{Z}^+$,

$$\int_0^1 (1-x^2)^n dx = \frac{2^{2n}(n!)^2}{(2n+1)^2}$$

Solution 1.1.1. Let $P(n)$ be the statement that $\displaystyle\int_0^1 (1-x^2)^n dx = \frac{2^{2n}(n!)^2}{(2n+1)^2}$.

- ⁎ The base case $P(1)$ is obviously true because

$$\int_0^1 (1-x^2)dx = \left[x - \frac{1}{3}x^3\right]_0^1 = \frac{2}{3} = \frac{2^2 \cdot 1^2}{3!}$$

- ⁎ Now we are at the inductive step. Assume $P(k)$ is true, i.e., $\displaystyle\int_0^1 (1-x^2)^k dx = \frac{2^{2k}(k!)^2}{(2k+1)^2}$. We will try to show $P(k+1)$ is also true. We can first integrate the left hand side (LHS hereafter and RHS for right hand side) of $P(k+1)$ by parts:

$$\int_0^1 (1-x)^{k+1} dx \qquad \begin{bmatrix} u = (1-x^2)^{k+1} & dv = dx \\ du = -2x(k+1)(1-x^2)^k dx & v = x \end{bmatrix}$$

$$\begin{aligned}
&= \left[x(1-x^2)^{k+1}\right]_0^1 + \int_0^1 2x^2(k+1)(1-x^2)^k dx \\
&= 2(k+1)\int_0^1 x^2(1-x^2)^k dx \\
&= 2(k+1)\int_0^1 [1-(1-x^2)](1-x^2)^k dx \\
&= 2(k+1)\int_0^1 (1-x^2)^k dx - 2(k+1)\int_0^1 (1-x^2)^{k+1} dx
\end{aligned}$$

Arranging like terms gives

$$\begin{aligned}
\int_0^1 (1-x^2)^{k+1} dx &= \frac{2k+2}{2k+3}\int_0^1 (1-x^2)^k dx \\
&= \frac{2k+2}{2k+3} \cdot \frac{2^{2k}(k!)^2}{(2k+1)^2} = \frac{(2k+2)\cdot 2^{2k}(k!)^2}{(2k+3)(2k+1)^2} \\
&= \frac{(2k+2)^2 \cdot 2^{2k}(k!)^2}{(2k+3)(2k+2)(2k+1)!} = \frac{4(k+1)^2 \cdot 2^{2k}(k!)^2}{(2k+3)!} \\
&= \frac{2^{2(k+1)}((k+1)!)^2}{(2(k+1)+1)^2}
\end{aligned}$$

and thus $P(k+1)$ true.

- Since $P(1)$ is correct and $P(k) \implies P(k+1) \forall k \in \mathbb{Z}$, the statement is true $\forall n \in \mathbb{Z}$. Thus,

$$\int_0^1 (1-x^2)^n dx = \frac{2^{2n}(n!)^2}{(2n+1)^2}$$

□

1.2 The Principle of Strong Induction

Theorem 1.2.1. Let a be a natural number, and let $P(n)$ be a statement about n for each integer $n \geqslant a$. If

(1) $P(a)$ is true, and

(2) for every integer $k \geqslant a$, if $P(n)$ holds for all $n \in [a, k]$, then $P(k+1)$ is also true,

then $P(k)$ is true for every integer $k \geqslant a$.

Problem 1.2.1 (1.3.9)**.** Let $\alpha = \dfrac{1+\sqrt{5}}{2}$ and $\beta = \dfrac{1-\sqrt{5}}{2}$, two roots of the quadratic equation $x^2 - x - 1 = 0$. Show that the n^{th} Fibonacci number satisfies

$$f_n = \frac{1}{\sqrt{5}}(\alpha^n - \beta^n)$$

Solution 1.2.1. Let $P(n)$ be the statement that $f_n = \dfrac{1}{\sqrt{5}}(\alpha^n - \beta^n)$.

It is clear that, starting from the third term, each Fibonacci number is determined by the sum of the two preceding terms, i.e., $f_n = f_{n-1} + f_{n-2}$ when $n \geqslant 3$. Therefore, the weak induction (1.1) would no longer apply. We need a proof involving $P(k-1) \wedge P(k) \implies P(k_1)$ and we need two base cases: $P(1)$ and $P(2)$.

- The base cases $P(1)$ and $P(2)$ are true because, by the equation given,

$$f_1 = \frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2} - \frac{1-\sqrt{5}}{2}\right) = 1$$

$$f_2 = \frac{1}{\sqrt{5}}\left(\left(\frac{1+\sqrt{5}}{2}\right)^2 - \left(\frac{1-\sqrt{5}}{2}\right)^2\right) = \frac{1}{\sqrt{5}}\left(\frac{(1+\sqrt{5})^2}{4} - \frac{(1-\sqrt{5})^2}{4}\right) = \frac{1}{\sqrt{5}} \cdot \frac{4\sqrt{5}}{4} = 1$$

  and f_1 and f_2 are indeed both 1.

- Now, for the inductive step, assume $P(n)$ holds for all positive integer not greater than k, among which are $P(k-1)$ and $P(k)$. So we know the following:

$$f_{k-1} = \frac{1}{\sqrt{5}}(\alpha^{k-1} - \beta^{k-1}) \text{ and } f_k = \frac{1}{\sqrt{5}}(\alpha^k - \beta^k)$$

  Since $f_{k+1} = f_{k-1} + f_k$, we have

$$\begin{aligned} f_{k+1} &= \frac{1}{\sqrt{5}}(\alpha^{k-1} - \beta^{k-1}) + \frac{1}{\sqrt{5}}(\alpha^k - \beta^k) \\ &= \frac{1}{\sqrt{5}}(\alpha^{k-1} + \alpha^k) - \frac{1}{\sqrt{5}}(\beta^{k-1} + \beta^k) \\ &= \frac{1}{\sqrt{5}}(\alpha^{k+1} - \beta^{k+1}) \text{ because } \alpha, \beta \text{ are roots of } x^2 = x + 1 \text{ and thus } x^{k+1} = x^k + x^{k-1}. \end{aligned}$$

- Now the base case and the inductive step are both proven, and we are done with the proof. □

Future reference: problem 2.1.1

Problem 1.2.2 (1.3.11)**.** Suppose that n is a positive integer. Prove, using strong induction, integration by parts, and l'Hôpital's Rule that

$$\int_0^1 (-\ln x)^n dx = n!$$

Solution 1.2.2. Let $P(n)$ be the statement that $\int_0^1 (-\ln x)^n dx = n!$. Clearly $P(0)$ is true, but I'll use $P(1)$ here as base case simply because I feel like anything raised to the 0^{th} power and 0! both feel somehow like exceptions.

- The base case $P(1)$ can be easily proven, but note that this is an improper integral because $\ln 0$ is undefined.

$$\int_0^1 (-\ln x)dx = \lim_{t\to 0^+}\int_t^1 (-\ln x)dx \qquad \begin{bmatrix} u = -\ln x & dv = dx \\ du = -\frac{1}{x} & v = x \end{bmatrix}$$

$$= \lim_{t\to 0^+}\left[[-x\ln x]_t^1 - \int_t^1 (-1)dx\right] = \lim_{t\to 0^+}\left[1 - t - [-x\ln x]_t^1\right]$$

$$= \lim_{t\to 0^+}[1 - t + t\ln t] = 1 + \lim_{t\to 0^+}[t\ln t] = 1 + \lim_{t\to 0^+}\left[\frac{\ln t}{\frac{1}{t}}\right]$$

$$\overset{(H)}{=} 1 + \lim_{t\to 0^+}\left[\frac{\frac{1}{t}}{-\frac{1}{t^2}}\right] = 1 + \lim_{t\to 0^+}(-t) = 1 = 1!$$

Thus the base case is correct.

- Now assume we have some positive integer k such that $P(n)$ is true $\forall n \in [1,k]$. We will try to integrate the LHS of $P(k+1)$:

$$\int_0^1 (-\ln x)^{k+1}dx = (-1)^{k+1}\int_0^1 (\ln x)^{k+1}dx \qquad \begin{bmatrix} u = (\ln x)^{k+1} & dv = dx \\ du = \frac{(k+1)(\ln x)^k}{x}dx & v = x \end{bmatrix}$$

$$= (-1)^{k+1}\lim_{t\to 0^+}\int_t^1 (\ln x)^{k+1}dx$$

$$= (-1)^{k+1}\lim_{t\to 0^+}\left[\left[t(\ln t)^{k+1}\right]_t^1 - \int_t^1 (k+1)(\ln x)^k dx\right]$$

$$= -(-1)(k+1)\left[(-1)^k \lim_{t\to 0^+}\int_t^1 (\ln x)^k dx\right] + (-1)^{k+1}\lim_{t\to 0^+}\left[t(\ln t)^{k+1}\right]_t^1$$

$$= (k+1)k! + (-1)^{k+1}\lim_{t\to 0^+}\left[t(\ln t)^{k+1}\right]_t^1 \text{ (by induction hypothesis)}$$

$$= (k+1)! + 0 - \lim_{t\to 0^+}\left[t(\ln t)^{k+1}\right]$$

We need to pause for a second and apply l'Hôpital's Rule $k+1$ times to evaluate the third term.

$$\begin{aligned}
\lim_{t\to 0^+}\left[t(\ln t)^{k+1}\right] &= \lim_{t\to 0^+}\left[\frac{(\ln t)^{k+1}}{\frac{1}{t}}\right] \\
&\overset{(\mathrm{H})}{=} \lim_{t\to 0^+}\left[\frac{\frac{(k+1)(\ln t)^k}{t}}{-\frac{1}{t^2}}\right] = -(k+1)\lim_{t\to 0^+}\left[t(\ln t)^k\right] = -(k+1)\lim_{t\to 0^+}\left[\frac{(\ln t)^k}{\frac{1}{t}}\right] \\
&\overset{(\mathrm{H})}{=} -(k+1)\lim_{t\to 0^+}\left[\frac{\frac{k(\ln t)^{k-1}}{t}}{-\frac{1}{t^2}}\right] = (k+1)k\lim_{t\to 0^+}\left[t(\ln t)^{k-1}\right] \\
&= \cdots \text{ (reiterate l'Hôpital's Rule for } k+1 \text{ times)} \\
&= (-1)^{k+1}(k+1)!\lim_{t\to 0^+}(t) = 0
\end{aligned}$$

Therefore,

$$\int_0^1 (-\ln x)^{k+1}dx = (k+1)!$$

and we have proven the inductive step.

⋆ Since the base case and the inductive step are proven, we are done with the proof. □

1.3 The Binomial Theorem

Definition 1.3.1 (Falling and rising factorials)**.** Just some convenient notations: the falling and rising factorials are defined in the two lines below, respectively.

$$\text{Rising factorial: } x^{\overline{n}} = x(x+1)(x+2)\cdots(x+n-1) = \prod_{k=1}^{n} x+k-1$$

$$\text{Falling factorial: } x^{\underline{n}} = x(x-1)(x-2)\cdots(x-n+1) = \prod_{k=1}^{n} x-k+1$$

Theorem 1.3.2 (Pascal's Identity)**.** If $0 < k < n+1$ then

$$\binom{n+1}{k} = \binom{n}{k-1} + \binom{n}{k}.$$

Proof. By definition,

$$\begin{aligned}\binom{n}{k+1} + \binom{n}{k} &= \frac{n!}{(k-1)!(n-k+1)!} + \frac{n!}{k!(n-k)!}\\ &= \frac{n!\cdot k + n!(n-k+1)}{k!(n-k+1)!}\\ &= \frac{(n+1)!}{k!(n+1-k)!} = \binom{n+1}{k}\end{aligned}$$

□

Theorem 1.3.3 (Binomial Theorem)**.** For any non-negative integer n and $x, y \in \mathbb{R}$ we have

$$(x+y)^n = \sum_{i=0}^{n} \binom{n}{i} x^{n-i} y^i$$

Proof. Let $P(n)$ be the statement that $(x+y)^n = \sum_{i=0}^{n} \binom{n}{i} x^{n-i} y^i$. We will prove this theorem by inducting on n.

- ⁎ One can immediately notice that the base case $P(0)$ is true:

$$(x+y)^0 = 1 = \binom{0}{0} x^0 y^0$$

- ⁎ Now, in the inductive step, we assume $P(k)$ is true, i.e., $(x+y)^k = \sum_{i=0}^{k} \binom{k}{i} x^{n-1} y^i$, and we want to show $P(k+1)$ is true. Multiplying the LHS of $P(k)$ with $(x+y)$ gives that of $P(k+1)$, and multiplying the RHS of $P(k)$ gives

$$\begin{aligned}(x+y)\sum_{i=0}^{k}\binom{k}{i}x^{k-i}y^i &= x\sum_{i=0}^{k}\binom{k}{i}x^{k-i}y^i + y\sum_{i=0}^{k}\binom{k}{i}x^{k-i}y^i\\ &= \sum_{i=0}^{k}\binom{k}{i}x^{k+1-i}y^i + \sum_{i=0}^{k}\binom{k}{i}x^{k-i}y^{i+1}\\ &= x^{k+1}y^0 + \left[\sum_{i=1}^{k}\binom{k}{i}x^{k+1-i}y^i + \sum_{i=0}^{k-1}\binom{k}{i}x^{k-i}y^{i+1}\right] + x^0y^{k+1}\\ &= x^{k+1}y^0 + \left[\sum_{i=1}^{k}\binom{k}{i}x^{k+1-i}y^i + \sum_{i=1}^{k}\binom{k}{i-1}x^{k+1-i}y^i\right] + x^0y^{k+1}\\ &= \binom{k+1}{0}x^{k+1}y^0 + \sum_{i=1}^{k}\binom{k+1}{i}x^{k+1-i}y^i + \binom{k+1}{k+1}x^0y^{k+1}\\ &= \sum_{i=0}^{k+1}\binom{k+1}{i}x^{k+1-i}y^i\end{aligned}$$

  which equals the RHS of $P(k+1)$ and proves the induction hypothesis.

- Now we've proven both the base case and the inductive step, and thus we've proven the Binomial Theorem.

  Future reference: problem 5.7.1

□

Problem 1.3.1 (1.4.5)**.** Prove that

$$\binom{r}{m}\binom{m}{k} = \binom{r}{k}\binom{r-k}{m-k}$$

Solution 1.3.1.

$$\begin{aligned}\binom{r}{m}\binom{m}{k} &= \frac{r!}{m!(r-m)!}\cdot\frac{m!}{k!(m-k)!}\\ &= \frac{r!}{k!(r-m)!(m-k)!} = \frac{r!(r-k)!}{k!(r-m)!(m-k)!}\\ &= \frac{r!}{k!(r-k)!}\cdot\frac{(r-k)!}{((r-k)-(m-k))!(m-k)!}\\ &= \binom{r}{k}\binom{r-k}{m-k}\end{aligned}$$

□

Problem 1.3.2 (1.4.12)**.** Prove the General Leibniz Rule:

$$(f\cdot g)^{(n)} = \sum_{i=0}^{n}\binom{n}{i}f^{(n-i)}g^{(i)}$$

where f and g are both functions of x.

Solution 1.3.2. Let $P(n)$ be the statement that $(f\cdot g)^{(n)} = \sum_{i=0}^{n}\binom{n}{i}f^{(n-i)}g^{(i)}$. We will approach this proof by (weak) induction.

- The base case is obviously true by the first-order product rule:

$$(f\cdot g)^{(1)} = f\cdot g^{(1)} + f^{(1)}\cdot g = \sum_{i=0}^{1}\binom{1}{i}f^{(1-i)}\cdot g^{(i)}$$

- For the inductive step, we want to show $P(k) \implies P(k+1)$. Taking derivatives of both sides of $P(k)$ should give us both sides of $P(k+1)$, respectively. Now suppose $(f\cdot g)^{(k)} = \sum_{i=0}^{k}\binom{k}{i}f^{(k-i)}\cdot g^{(i)}$, then

$$\begin{aligned}
(f\cdot g)^{(k+1)} &= \frac{d}{dx}\left[\sum_{i=0}^{k}\binom{k}{i}f^{(k-i)}\cdot g^{(i)}\right]\\
&=\sum_{i=0}^{k}\binom{k}{i}f^{(k-i)}\cdot g^{(i+1)}+\sum_{i=0}^{k}\binom{k}{i}f^{(k-i+1)}\cdot g^{(i)}\\
&=\sum_{i=1}^{k+1}\binom{k}{i-1}f^{(k-i+1)}\cdot g^{(i)}+\sum_{i=0}^{k}\binom{k}{i}f^{(k-i+1)}\cdot g^{(i)}\\
&=\left[\binom{k}{(k+1)-1}f^{(k-(k+1)+1)}\cdot g^{(k+1)}+\sum_{i=1}^{k}\binom{k}{i-1}f^{(k-i+1)}\cdot g^{(i)}\right]\\
&\quad+\left[\sum_{i=1}^{k}\binom{k}{i}f^{(k-i+1)}\cdot g^{(i)}+\binom{k}{0}f^{(k+1)}\cdot g^{(0)}\right]\\
&=f\cdot g^{(k+1)}+\left[\sum_{i=1}^{k}\binom{k}{i-1}f^{(k-i+1)}\cdot g^{(i)}+\sum_{i=1}^{k}\binom{k}{i}f^{(k-i+1)}\cdot g^{(i)}\right]+f^{(k+1)}\cdot g\\
&=f\cdot g^{(k+1)}+\sum_{i=1}^{k}\left[\binom{k}{i-1}+\binom{k}{i}\right]f^{(k-i+1)}\cdot g^{(i)}+f^{(k+1)}\cdot g\\
&=f\cdot g^{(k+1)}+\sum_{i=1}^{k}\binom{k+1}{i}f^{(k-i+1)}\cdot g^{(i)}+f^{(k+1)}\cdot g\\
&=\sum_{i=0}^{k+1}\binom{k+1}{i}f^{((k+1)-i)}\cdot g^{(i)}
\end{aligned}$$

hence proven the inductive step.

- ⋆ With the base case and inductive step both proven, we have just proven the General Leibniz Rule. □

Problem 1.3.3 (1.4.16, Vandermonde's Identity)**.** Prove that, given $a \geqslant n \geqslant 0$ and $b \geqslant n \geqslant 0$,

$$\sum_{i=0}^{n}\binom{a}{i}\binom{b}{n-i}=\binom{a+b}{n}$$

Solution 1.3.3–1. Different from before, this time we will use induction on b to approach this proof. Let $P(b)$ be the statement that $\sum\limits_{i=0}^{n}\binom{a}{i}\binom{b}{n-i}=\binom{a+b}{n}$.

- ⋆ The base case $P(0)$ is clearly true because

$$\sum_{i=0}^{n}\binom{a}{i}\binom{0}{n-i}=\sum_{i=0}^{n-1}\binom{a}{i}\binom{0}{n-i}+\binom{a}{n}=\binom{a}{n}$$

since $\binom{x}{y}=0$ when $x<y$ by definition.

- ⋆ Now, for the inductive step, assume that $P(k)$ holds, i.e., $\sum\limits_{i=0}^{n}\binom{a}{i}\binom{k}{n-i}=\binom{a+k}{n}$. We want to show that $P(k+1)$ is also true. We will approach this by applying the induction hypothesis and

Pascal's Identity.

$$\begin{aligned}
\binom{a+(k+1)}{n} &= \binom{a+k}{n} + \binom{a+k}{n-1} && \text{(Pascal's Identity)}\\
&= \sum_{i=0}^{n}\binom{a}{i}\binom{k}{n-i} + \sum_{i=0}^{n-1}\binom{a}{i}\binom{k}{n-1-i} && (P(k))\\
&= \binom{a}{n}\binom{k}{0} + \sum_{i=0}^{n-1}\binom{a}{i}\binom{k}{n-i} + \sum_{i=0}^{n-1}\binom{a}{i}\binom{k}{n-1-i}\\
&= \binom{a}{n}\binom{k+1}{0} + \sum_{i=0}^{n-1}\binom{a}{i}\left[\binom{k}{n-i} + \binom{k}{n-1-i}\right]\\
&= \binom{a}{n}\binom{k+1}{0} + \sum_{i=0}^{n-1}\binom{a}{i}\binom{k+1}{n-i} && \text{(Pascal's Identity)}\\
&= \sum_{i=0}^{n}\binom{a}{i}\binom{k+1}{n-i}
\end{aligned}$$

This concludes the proof of the inductive step.

⋆ We have proven the base case, the inductive step, and thus Vandermonde's Identity. □

Solution 1.3.3–2. Consider the binomial expansion of $(1+x)^{a+b}$. By the Binomial Theorem, the coefficient of x^n is $\binom{a+b}{n}$.

Also notice that $(1+x)^{a+b} = (1+x)^a \cdot (1+x)^b$, and we can apply the Binomial Theorem to both terms separately.

$$(1+x)^{a+b} = (1+x)^a \cdot (1+x)^b = \left(\sum_{i=0}^{a}\binom{a}{i}x^i\right)\left(\sum_{j=0}^{b}\binom{b}{j}x^j\right)$$

Now if we try to find the coefficient of x^n again, we can start by checking the case where $i = 0$ and $j = n$, then $i = 1$ and $j = n-1$, and so on, until $x = n$ and $j = 0$. Each case has coefficient $\binom{a}{i}\binom{b}{n-i}$, and we know their sum is $\binom{a+b}{n}$ as previously shown. Therefore,

$$\sum_{i=0}^{n}\binom{a}{i}\binom{b}{n-i} = \binom{a+b}{n}$$

hence proven. □

2 Arithmetic

2.1 Divisibility and the GCD

Problem 2.1.1. [2.1.6] Let f_n denoteh the n^{th} Fibonacci number. Prove, using induction, that $\gcd(f_n, f_{n+1}) = 1$ for all $n \in \mathbb{Z}^+$.

Solution 2.1.1. As mentioned in solution for problem 1.2.1, our base case involves two terms, f_1 and f_2.

- The base case is true because $f_1 = f_2 = 1$ and $\gcd(f_1, f_2) = \gcd(1,1) = 1$.
- Now assume $\gcd(f_k, f_{k+1}) = 1$ is a true statement. Then $\gcd((f_k + f_{k+1}), f_{k+1}) = 1$ is also true. Since $f_{k+2} = f_k + f_{k+1}$, our inductive step is also proven.
- Therefore, each pair of consecutive Fibonacci numbers is co-prime.

2.2 The Division Algorithm

Problem 2.2.1 (2.2.3)**.** Prove that if n is a perfect square then n musc have the form $4k$ or $4k+1$ where $k \in \mathbb{N}$.

Solution 2.2.1. Since n is a perfect square, $n = m^2$ for some $m \in \mathbb{N}>$

- If $m \equiv 0 \pmod 4$, then $\exists p \in \mathbb{N}$ such that $m = 4p$ and $n = 16p^2 = 4(4p^2)$.
- If $m \equiv 1 \pmod 4$, then $\exists p \in \mathbb{N}$ such that $m = 4p+1$ and $n = 16p^2 + 8p + 1 = 4(4p^2 + 2p) + 1$.
- If $m \equiv 2 \pmod 4$, then $\exists p \in \mathbb{N}$ such that $m = 4p+2$ and $n = 16p^2 + 16p + 4 = 4(4p^2 + 4p + 1)$.
- If $m \equiv 3 \pmod 4$, then $\exists p \in \mathbb{N}$ such that $m = 4p+3$ and $n = 16p^2 + 24p + 9 = 4(4p^2 + 6p + 2) + 1$.

We see n always has the form $4k$ and $4k+1$ regardless of its residue class. □

2.3 Euclid's Algorithm

Definition 2.3.1 (Euclid's Algorithm)**.** Given $x > y$ and $x, y \in \mathbb{Z}^+$, the following process is called Euclid's Algorithm:

$$\begin{aligned} x &= yq_1 + r_1 && 0 \leqslant r_1 < b \\ b &= r_1q_2 + r_2 && 0 \leqslant r_2 < r_1 \\ r_1 &= r_2q_3 + r_3 && 0 \leqslant r_3 < r_2 \\ &\vdots && \vdots \\ r_{i-2} &= r_{i-1}q_i + r_i && 0 \leqslant r_i < r_{i-1} \\ r_{i-1} &= r_iq_{i+1} + r_{i+1} && r_{i+1} = 0 \end{aligned}$$

Since the sequence (r_n) is strictly decreasing, it will eventually reach 0, as shown by the last line above. The last non-zero remainder, r_i, satisfies $r_i = \gcd(x, y)$.

Future reference: theorem 2.4.1, Euclid's Algorithm for Polynomials

2.4 ℤ Combination of Two Integers

Theorem 2.4.1. Suppose $x, y \in \mathbb{Z}$ are not *both* zero, then $\exists a, b \in \mathbb{Z}$ such that

$$ax + by = \gcd(x, y)$$

Proof. We will first show that this statement holds true $\forall x, y \in \mathbb{Z}^+$. Without loss of generality (WLOG hereafter), assume $a > b$. Recall the remainder r_n in Euclid's Algorithm. It follows that r_1 is a $\mathbb{Z}$ combination of x and y since $r_1 = x - yq_1$. Then, r_{i+1} can always be obtained by a $\mathbb{Z}$ combination of r_{i-1} and q_i since $r_{i+1} = r_{i-1} - r_i q_{i+1}$. Thus r_n is always obtainable using a $\mathbb{Z}$ combination of x and y, and the statement holds $\forall x, y \in \mathbb{Z}^+$.

If $x \in \mathbb{Z}^-$ and $y \in \mathbb{Z}^+$ then $-x \in \mathbb{Z}^+$ and from the case above we know $\exists a_1, b_1 \in \mathbb{Z}$ such that

$$a_1(-x) + b_1 y = \gcd(-x, y).$$

Since $\gcd(-x, y) = \gcd(x, y)$ and $a_1(-x) = (-a_1)x$, we have

$$(-a_1)x + b_1 y = \gcd(x, y)$$

Likewise if $x \in \mathbb{Z}^+$ and $y \in \mathbb{Z}^-$.

If $x = 0$ and $y \in \mathbb{Z}^+$ then $\gcd(x, y) = y = 0 \cdot x + 1 \cdot y$.

If $x = 0$ and $y \in \mathbb{Z}^-$ then $\gcd(x, y) = y = 0 \cdot x + (-1) \cdot y$.

Likewise for the two cases given $y = 0$.

Future reference: Bézout's Identity, theorem 4.2.1, theorem 5.3.8 □

Theorem 2.4.2 (Bézout's Identity)**.** Let x and y be integers with $\gcd(x, y) = d$. Then $\exists a, b \in \mathbb{Z}$ such that $ax + by = d$. More generally, all integers of the form $ax + by$ are multiples of d.

Proof. The first part is already proven in theorem 2.4.1. The second part is more obvious. Since $\gcd(x, y) = d$, we have $d \mid x$ and $d \mid y$. It follows that any $\mathbb{Z}$ combination of two multiples of d will still be a multiple of d, and we are done with the proof. □

2.5 The Fundamental Theorem of Arithmetic

Theorem 2.5.1 (Fundamental Theorem of Arithmetic, FTA)**.** Each positive integer greater than 1 can be uniquely factorized.

Proof. Suppose not, then there exists some positive integer x such that it can be prime factorized in at least two different ways:

$$x = \prod_{i=1}^{k} p_i \text{ and } x = \prod_{i=1}^{\ell} q_i \text{ where } p_i < p_{i+1}$$

WLOG, assume $k \leqslant \ell$. Since $p_1 \mid x$, there exists some q_m such that $p_1 \mid q_i$. However, to divide q_m, p_1 has to either equal to 1 or q_m, but any prime number is greater than 1. Therefore $p_1 = q_m$. Rearrange the indexes of the rest of q's so that they are now named $q_2, q_3, \cdots, q_j$. Similar to $p_1 = q_m$, there exists another q_n such that $p_2 = q_n$. Since $k \leqslant \ell$, each prime number in the p-prime factorization is equal to some prime number in the q-factorization. With proper rearrangement, we have shown that

$$\prod_{i=1}^{k} p_i = \prod_{i=1}^{k} q_i$$

Therefore,

$$\frac{x}{\prod_{i=1}^{k} p_i} = \frac{x}{\prod_{i=1}^{k} q_i} \implies 1 = \prod_{i=k+1}^{\ell} q_i$$

which is clearly impossible, given that each prime number is greater than 1. Therefore, the assumption that there exists a positive integer greater than 1 that can be prime factorized in more than one way is false, and we've proven the FTA. □

Theorem 2.5.2. There are infinite prime numbers.

Proof. Suppose not, then we can list all the prime numbers, which we call $p_1, p_2, \cdots, p_n$. Then consider the number $x = 1 + \prod\limits_{i=1}^{n} p_i$. If x is prime then we immediately have a contradiction that suggests prime numbers are infinite. If not, then it must be a multiple of some p_k that's on the list of "all" prime numbers. Now consider the following equation:

$$x - \prod_{i=1}^{n} p_i = 1$$

Since both x and $\prod\limits_{i=1}^{n} p_i$ are multiples of p_k, it follows that $p_k \mid 1$ which is clearly a contradiction. Therefore the assumption that prime numbers are finite must be false and we are done with the proof. □

3 Complex Numbers, Sets, and Functions

3.1 Complex Numbers

Some basic concepts...

- $\Re(a+bi)=a; \Im(a+bi)=b; |a+bi|=\sqrt{a^2+b^2}$
- $\Re(z), \Im(z), |z| \in \mathbb{R}$
- $(a+bi)+(c+di)=(a+c)+(b+d)i$
- $(a+bi)(c+di)=(ac-bd)+(ad+bc)i$
- Conjugate: $z=a+bi \implies \overline{z}=a-bi$
- $z_1 \cdot z_2 = z_2 \cdot z_1$
- $|z_1 \cdot z_2| = |z_1| \cdot |z_2|$
- $\overline{z_1 \cdot z_2} = \overline{z_1} \cdot \overline{z_2}$
- $z \cdot \overline{z} = |z|^2$
- Polar form: $r(\cos\theta + i\sin\theta)$

Some important formulas:

Proposition 3.1.1 (Euler's Formula)**.**

$$e^{i\theta} = \cos\theta + i\theta \quad \forall\theta \in \mathbb{R}$$

Proposition 3.1.2 (DeMoivre's Theorem)**.** Given $z \in \mathbb{C}$ in polar form,

$$z = r(\cos\theta + i\sin\theta) \iff z^n = r^n(\cos(n\theta) + i\sin(n\theta))$$

Future reference: theorem 3.2.2, problem 5.5.1

3.2 Roots of Unity

Definition 3.2.1. Given $n \in \mathbb{Z}^+$, an n^{th} root of unity is a $z \in \mathbb{C}$ such that $z^n = 1$. The set of all n^{th} roots of unity is denoted as

$$\mu_n = \{z \in \mathbb{C} : z^n = 1\}.$$

Theorem 3.2.2. Suppose $n \in \mathbb{Z}^+$ and let $\zeta = e^{2\pi i/n}$. Then

$$\mu_n = \left\{\zeta^k : k \in \mathbb{Z}\right\}.$$

Proof. To prove the equality, it suffices to show that $\mu_n \subseteq \left\{\zeta^k : k \in \mathbb{Z}\right\}$ and $\left\{\zeta^k : k \in \mathbb{Z}\right\} \subseteq \mu_n$, respectively. The second statement is obvious. Since ζ is given as an n^{th} root of unity, we already know $\zeta^n = 1$. All we need to show is that $\forall k \in \mathbb{Z}, \zeta^k \in \mu_n$, i.e., $(\zeta^k)^n = 1$. A tiny amount of rearrangement gives

$$(\zeta^k)^n = \zeta^{kn} = (\zeta^n)^k = 1^k = 1$$

which proves that $\left\{\zeta^k : k \in \mathbb{Z}\right\} \subseteq \mu_n$. Now look at the first statement. By definition, we know that $\forall z \in \mu_n, z^n = 1$. If we express z in polar form as $re^{i\theta}$, then by DeMoivre's Theorem,

$$1 = z^n = r^n e^{in\theta}$$

which immediately suggests $r = 1$ and $n\theta = 2\pi k$ for some $k \in \mathbb{Z}$. This is exactly the same k that we are looking for since

$$z = re^{i\theta} = e^{2\pi ki/n} = \left(e^{2\pi i/n}\right)^k = \zeta^k$$

which proves that $\mu_n \subseteq \left\{\zeta^k : k \in \mathbb{Z}\right\}$. Therefore $\mu_n = \left\{\zeta^k : k \in \mathbb{Z}\right\}$. □

Definition 3.2.3. A complex number z has order d if $z^d = 1$ and $z \neq 1$ for all $0 \leqslant m < d$. A ***primitive*** n^{th} ***root of unity*** is a $z \in \mu_n$ of order n. A primitive n^{th} root of unity is a $z \in \mu_n$ of order n.

(Example: $z_1 = e^{\pi i/5} = e^{2\pi i/10}$ is a primitive 10^{th} root of unity but $z_2 = e^{2\pi i/5} = e^{4\pi i/10}$ isn't because $(z_2)^5 = e^{2\pi i} = 1$.)

Theorem 3.2.4. Suppose z is a root of unity of order d. Then for any $k \in \mathbb{Z}$,
z^k has order $d/\gcd(d,k)$

Proof. Let $d_0 = d/\gcd(d,k)$ and $k_0 = k/\gcd(d,k)$. It follows that

$$\gcd\left(\frac{d}{\gcd(d,k)}, \frac{k}{\gcd(d,k)}\right) = \gcd(d_0, k_0) = 1.$$

Clearly, $kd_0 = \dfrac{kd}{\gcd(d,k)} = dk_0$. Therefore,

$$(z^k)^{d_0} = z^{kd_0} = z^{dk_0} = (z^d)^{k_0} = 1$$

and z^k is a $(d_0)^{th}$ root of unity. Now we need to show that z^k is a primitive $(d_0)^{th}$ root of unity, i.e., no positive integer smaller $f < d_0$ satisfy $(z^k)^f = 1$.

Suppose there exists such $f \in \mathbb{Z}^+$ such that $f < d_0$ and $(z^k)^f = 1$. Then since $z^{(kf)} = 1$, it follows that kf is a multiple of d. Dividing both kf and d by $\gcd(d,k)$ we have $d_0 \mid k_0 f$. However, we know

that $\gcd(d_0, k_0) = 1$, so $d_0 \mid f$, which suggests $d_0 \leqslant f$, contradiction. Therefore if z is a root of unity of order d, z^k is a root of unity of order $d/\gcd(d, k)$.
Future reference: problem 3.2.1, prob 4.1.1 □

Problem 3.2.1 (3.2.18, old version)**.** Suppose p is a prime number. Show that there are $p^n - p^{n-1}$ primitive p^n-th roots of unity.

Solution 3.2.1. Let $\zeta = e^{2\pi i/p^n}$. Clearly ζ is a primitive p^n-th root of unity. The set of all p^n-th roots of unity has p^n distinct elements:

$$\mu_{p^n} = \left\{\zeta^k : 1 \leqslant k \leqslant p^n\right\}$$

However, not all of these elements are primitive p^n-th root of unity. By theorem 3.2.2, ζ^k is a primitive p^n-th root of unity if and only if k and p^n are co-prime. Since p is a prime number, the only case when k and p^n aren't co-prime is when $p \mid k$. There are $p^n/p = p^{n-1}$ such numbers between 1 and p^n. Excluding these non-primitive roots of unity, we therefore have $p^n - p^{n-1}$ primitive p^n-th roots of unity.

Future reference: 4.4.1

3.3 Operations on Sets

Proposition 3.3.1 (De Morgan's Laws)**.** For any sets A and B,

$$(A \cup B)^C = A^C \cap B^C$$
$$(A \cap B)^C = A^C \cup B^C$$

Proposition 3.3.2 (Distributive laws)**.** For any sets A, B, and C, we have

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$
$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

3.4 Functions

Definition 3.4.1. Let $f : X \to Y$ be a function.

- f is ***injective*** if for every $x_1, x_2 \in X$,

  $$x_1 \neq x_2 \implies f(x_1) \neq f(x_2)$$

  or, rephrased by the contrapositive,

  $$f(x_1) = f(x_2) \implies x_1 = x_2$$

- f is ***surjective*** if for every $y \in Y$ there exists some $x \in X$ such that $f(x) = y$.
- f is ***bijective*** if it is injective and surjective'

Definition 3.4.2. For any set X define the ***identity function*** $\mathrm{id}_X : X \to X$ by $\mathrm{id}_X(x) = x$.

Definition 3.4.3. Suppose $f : X \to Y$ is a function. A function $g : Y \to X$ is called an ***inverse*** of f it it satisfies both relations

$$f \circ g = \mathrm{id}_Y \quad g \circ f = \mathrm{id}_X$$

Or, equivalently,

$$f(f^{-1}(y)) = y \ \forall y \in Y \text{ and } f^{-1}(f(x)) = x \ \forall x \in X$$

Future reference: theorem 3.4.4

Theorem 3.4.4. A function $f : X \to Y$ is invertible if and only if it is bijective.

Proof. For the $\implies$ it suffices to show that if f is invertible then it is both injective and surjective.

Suppose f is invertible then there exists an $f^{-1} : Y \to X$. Then,

$$\begin{aligned} f(x_1) = f(x_2) &\implies f^{-1}(f(x_1)) = f^{-1}(f(x_2)) && \text{(Definition of a function)} \\ &\implies x_1 = x_2 && \text{(Property of inverse function)} \end{aligned}$$

which shows f is injective. Now, because we know

$$f(f^{-1}(y)) = y \ \forall y \in Y$$

from definition 3.4.3, f is indeed surjective.

Now for the $\Longleftarrow$ direction, assume that f is bijective. Then for every $y \in Y$, there is an unique x such that $f(x) = y$. Now define $g : Y \to X$ as the following:

$$\text{For every } y \in Y\text{, let } g(y) \in X \text{ be the unique element such that } f(g(y)) = y$$

It follows that $\forall y \in Y, f(g(y)) = y$. Therefore $f \circ g = \text{id}_Y$. Now combining the two functions, $f(x)$ and $f(g(y))$, we have

$$\begin{cases} f(x) = y \\ f(g(y)) = y \end{cases} \implies g(y) = x \implies g(f(x)) = x \implies g \circ f = \text{id}_X$$

By definition 3.4.3, we've shown g is an inverse of f. $\Longleftarrow$ direction done as well. □

Problem 3.4.1 (3.4.14)**.** Suppose we have functions $f : A \to B$ and $g : B \to C$.

(a) Prove that if f and g are both injective then so is $g \circ f$.

(b) Prove that if f and g are both surjective then so is $g \circ f$.

(c) It follows from the previous two parts that if f and g are bijective then so is $g \circ f$. Is the converse

$$g \circ f \text{ bijective} \implies f \text{ and } g \text{ bijective}$$

true? Prove or give a counterexample.

Solution 3.4.1. Note that $g \circ f$ is a composite function that maps elements from A to C.

1. By definition of injecrtivity, $g(f(a)) = g(f(b)) \implies f(a) = f(b)$ since g is injective, and $f(a) = f(b) \implies a = b$ since f is injective. Therefore

$$(g \circ f)(a) = (g \circ f)(b) \iff g(f(a)) = g(f(b)) \implies a = b$$

hence $g \circ f$ is injective.

2. Since g is surjective, it follows that for each $c \in C$, there exists some $b \in B$ such that $g(b) = c$. But we also know f is surjective, and by definition there also exists some $a \in A$ such that $f(a) = b$. Therefore $\forall c \in C, \exists a \in A$ such that $(g \circ f)(a) = g(f(a)) = c$. Hence we've shown that $g \circ f$ is surjective.

3. False. Consider the functions

$$\begin{cases} f : \mathbb{R}^+ \to \mathbb{R}, & f(x) = x \\ g : \mathbb{R} \to \mathbb{R}^+, & g(x) = |x| \end{cases} \implies g \circ f : \mathbb{R}^+ \to \mathbb{R}^+, (g \circ f)(x) = |x|$$

in which case f is NOT surjective, g is NOT injective, but $g \circ f$ is bijective.

Problem 3.4.2 (3.4.19)**.** Define a function $f : \mathbb{Z}^+ \times \mathbb{Z}^+ \to \mathbb{Z}^+$ by the formula

$$f(a,b) = \begin{cases} \binom{a}{b} & \text{if a} \geqslant b \\ \binom{b}{a} & \text{if } b < a \end{cases}$$

Determine if f is injective, surjective, both, or neither.

Solution 3.4.2. Suppose $a \leqslant b$, we have $\binom{a}{b} = \binom{a}{a-b}$, which implies $f(a,b) = f(a,a-b)$ even though the ordered pairs $\langle a,b\rangle$ and $\langle a,a-b\rangle$ are not necessarily equal. Therefore f is not injective.

Notice that $\forall a \in \mathbb{Z}^+$, we always have $\binom{a}{1} = a$. Therefore, it we fix $b = 1$, $f(a,b)$ surjects the entire $\mathbb{Z}^+$. Thus, f is surjective.

Problem 3.4.3 (3.4.20)**.** Define a function $f : \mathbb{Z}^+ \times \mathbb{Z}^+ \to \mathbb{Z}^+$ by the formula

$$f(a,b) = \binom{a+b}{b}.$$

Determine if f is injective, surjective, both, or neither.

Solution 3.4.3. Similar to the previous question, since $\binom{a+b}{a} = \binom{a+b}{b}$, we have $f(a,b) = f(b,a)$ even though the ordered pairs $\langle a,b\rangle$ and $\langle b,a\rangle$ are not necessarily equal. Therefore f is not injective.

In this problem, f is not surjective either because the only occasions when $\binom{a+b}{b} = 0$ is $a+b=b$ or $b = 0$, both of which are impossible given $a, b \in \mathbb{Z}+$. Therefore the image of f does not include 1, and f is neither injective nor bijective.

3.5 Image and Preimage

Definition 3.5.1. Suppose $f : X \to Y$ is a function and $A \subseteq X$, $B \subseteq Y$.

(1) The ***image*** of A under f is the set

$$f(A) = \{f(a) : a \in A\}$$

(2) The ***preimage*** of B under f is the set

$$f^{-1}(B) = \{x \in X : f(x) \in B\}$$

Problem 3.5.1 (3.5.5-8)**.**

1. Suppose $B \subset Y$. Prove or give a counterexample to each of the inclusions

$$B \subset f(f^{-1}(B)) \qquad f(f^{-1}(B)) \subset B.$$

2. Suppose f is surjective, and $B \subset Y$. Prove that $f(f^{-1}(B)) = B$.

3. Suppose $A \subset X$. Prove or give a counterexample to each of the inclusions

$$A \subset f^{-1}(f(A)) \qquad f^{-1}(f(A)) \subset A.$$

4. Suppose f is injective, and $A \subset X$. Prove that $f^{-1}(f(A)) = A$.

Solution 3.5.1.

1. The first inclusion is false and the second is true.

   For the first one, consider $X = \{0\}$, $Y = \{0,1\}$, $B = \{1\}$, and the identity function $\mathrm{id}_X : X \to Y$. Then, $f^{-1}(B) = \varnothing$ since no $x \in X$ satisfy $f(x) = 1$. It follows that $f(f^{-1}(B)) = f(\varnothing) = \varnothing$, and $B \not\subset \varnothing$. Hence we've found a counterexample.

   The other inclusion, however, is correct. By definition, for all element $x \in f^{-1}(B)$, we always have $f(x) \in B$. Therefore $f(f^{-1}(B)) \subset B$. □

2. We have shown that $f(f^{-1}(B)) \subset B$ regardless of the surjectivity of f. Therefore it suffices to show that f is surjective $\implies B \subset f(f^{-1}(B))$.

   Choose an arbitrary $b \in B$. Since f is surjective, there exists $x \in X$ such that $f(x) = b$. Thus $x \in f^{-1}(B)$. Substituting b by $f(x)$, we have $b = f(x) \in f(f^{-1}(B))$. Since b is chosen arbitrarily, $B \subset f(f^{-1}(B))$ and thus $f(f^{-1}(B)) = B$. □

3. The first inclusion is true and the second is false.

   The first inclusion always holds. For $a \in A$, it is always true that $a \in f^{-1}(f(A))$ since $f(a) = f(a)$.

   As a counterexample to the second inclusion, consider $X = \{0,1\}$,$A = \{0\}$, $Y = \{0\}$, and $f : X \to Y$ defined by $f(0) = f(1) = 0$. Then, $f(A) = \{0\}$, and $f^{-1}(0) = \{0,1\}$. Clearly $\{0,1\} \not\subset \{0\}$. Hence this is a counterexample.

4. We have shown that $f^{-1}(f(A)) \subset A$ regardless of the injectivity of f. Therefore it suffices to show that f is injective $\implies f^{-1}(f(A)) \subset A$.

   Choose an arbitrary $x \in f^{-1}(f(A))$. By definition $f(x) \in f(A)$. Therefore there exists $a \in A$ such that $f(a) = f(x)$. Since f is injective, by (contrapositive) definition we know $a = x$. Hence if $x \in f^{-1}(f(A))$ then $x \in A$, i.e., $f^{-1}(f(A)) \subset A$. Therefore if f is injective then $f^{-1}(f(A)) = A$. □

Problem 3.5.2 (old version only)**.**

(a) Suppose that $A, B \subset X$, and that f is injective. Prove that

$$f(A) \subset f(B) \implies A \subset B.$$

(b) Show that the previous claim is false if we omit the hypothesis that f is injective.

(c) Suppose that $C, D \subset Y$, and that f is surjective. Prove that

$$f^{-1}(C) \subset f^{-1}(D) \implies C \subset D.$$

(d) Show that the previous claim is false if we omit the hypothesis that f is surjective.

Solution 3.5.2.

(a) Suppose $f(A) \subset f(B)$. Choose arbitrary $a \in A$. We immediately know $f(a) \in f(A)$ and $f(a) \in f(B)$ as well since $f(A) \subset f(B)$. By definition, given $f(a) \in f(B)$, there exists $b \in B$ such that $f(b) = f(a)$. However, because f is injective, $f(b) = f(a)$ implies $b = a$ and thus $a \in B$. Therefore if $f(A) \subset f(B)$ then $A \subset B$. □

(b) Counterexample: consider $A = X = \{0,1\}$, $B = \{0\}$, $Y = \{0\}$, and $f : X \to Y$ defined by $f(0) = f(1) = 0$. Then we have $f(A) = \{0\}$ and $f(B) = \{0\}$. Thus $f(A) \subseteq f(B)$. However, $A = \{0,1\} \not\subset \{0\} = B$.

(c) Suppose $f^{-1}(C) \subset f^{-1}(D)$. Choose an arbitrary $c \in C$. Since f is surjective, there exists $x \in X$ such that $f(x) = c$. Therefore $x \in f^{-1}(C)$. Since $f^{-1}(C) \subset f^{-1}(D)$, it follows that $x \in f^{-1}(D)$ as well. Then, again, by definition, $f(x) = c \in D$. Hence $C \subset D$. □

(d) Counterexample: consider $X = \{0\}$, $C = Y = \{0, 1\}$, $D = \{0\}$, and $\text{id}_X : X \to Y$. Then $f^{-1}(C) = \{0\}$ and $f^{-1}(D) = \{0\}$. We have $f^{-1}(C) \subseteq f^{-1}(D)$ but clearly $C = \{0, 1\} \not\subset \{0\} = D$.

Problem 3.5.3 (3.5.9)**.** Suppose $f^{-1}(f(A)) = A$ holds for <u>every</u> $A \subset X$. Prove that f is an injection.

Solution 3.5.3. Again, showing $f^{-1}(f(A)) = A$ is equivalent to showing both $f^{-1}(f(A)) \subset A$ and $A \subset f^{-1}(f(A))$. Since we've shown that the latter is true regardless of the injectivity of f, it suffices to show $f^{-1}(f(A)) \subset A \implies f$ is injective.

Suppose f is not injective. Then there exist $x_1, x_2 \in X$ such that $x_1 \neq x_2$ but $f(x_1) = f(x_2)$. If we let $A = \{x_1\}$ then we have $\{x_1, x_2\} = f^{-1}(f(A))$ but $\{x_1, x_2\} \not\subset A$. Contradiction. (Likewise if we let $A = x_2$.) Hence if $f^{-1}(f(A)) = A$ holds for every $A \subset X$ then f is injective. □

4 Congruences and the Ring $\mathbb{Z}/n\mathbb{Z}$

4.1 Equivalence Relations and Partitions

Definition 4.1.1. Let X be a set and let R be a relation on X. We say that R is an ***equivalence relation*** if it satisfies the following three properties:

(1) Reflexivity: every $a \in X$ satisfies aRa.

(2) Symmetry: for $a, b \in X$, aRb implies bRa.

(3) Transitivity: for $a, b, c \in X$, $aRb \wedge bRc \implies aRc$.

Definition 4.1.2. Let X be a set and let R be an equivalence relation on X. For any $a \in X$ define the equivalence class of a by

$$[a]_R = \{x \in X : aRx\}$$

Definition 4.1.3. Fix a positive integer n. The ***congruence modulo*** n is defined as

$$a \equiv b \pmod{n} \iff n \mid a - b$$

Definition 4.1.4. Let X be any set. A ***partition*** of X is a collection $\mathcal{C}$ of subsets of X with the following properties:

(1) For every $x \in X$ there is a $B \in \mathcal{C}$ such that $x \in B$

(2) Every $B \in \mathcal{C}$ is nonempty.

(3) For sets $B, B' \in \mathcal{C}$, either $B = B'$ or $B \cap B' = \varnothing$.

The elements $B \in \mathcal{C}$ are called blocks of partition.

Example 4.1.1. $\mathcal{C} = \{[0]_3, [1]_3, [2]_3\}$ is a partition of $\mathbb{Z}$ where $[x]_3$ denotes the equivalence class of x modulo 3.

(1) Each integer belongs to either $[0]_3, [1]_3$, or $[2]_3$.

(2) None of the blocks are empty.

(3) Clearly $[1]_3, [2]_3, [3]_3$ are different, and the intersection between any two of the three is $\varnothing$.

Problem 4.1.1 (4.1.16). Define a relation $\sim$ on μ_{10} by

$$z_1 \sim z_2 \iff (\text{the order of } z_1^4) = (\text{the order of } z_2^4)$$

Verify that $\sim$ is an equivalence relation, and determine the associated partition of μ_{10}.

Solution 4.1.1. Clearly $\sim$ satisfies reflexivity as z_1^4 always has the same order as itself. It satisfies symmetry because the symbol = is reflexive. It also satisfies transitivity because the symbol = is transitive.

Now let's determine the partition. Let $\zeta = e^{2\pi i/10}$. Then

$$\mu_{10} = \left\{\zeta, \zeta^2, \cdots, \zeta^{10}\right\}$$

Below is a diagram that shows the order of each $\zeta^k \in \mu_{10}$ as well as that of $(\zeta^k)^4$. Recall theorem 3.2.4 that if z is a root of unity of order d then z^k has order $d/\gcd(d,k)$.

Roots	ζ	ζ^2	ζ^3	ζ^4	ζ^5	ζ^6	ζ^7	ζ^8	ζ^9	ζ^{10}
Order of ζ^k	10	5	10	5	2	5	10	5	10	1
Order of $(\zeta^k)^4$	5	5	5	5	1	5	5	5	5	1

Therefore, the partition for this question would be

$$\mathcal{C} = \left\{\left\{\zeta^5, \zeta^{10}\right\}, \left\{\zeta, \zeta^2, \zeta^3, \zeta^4, \zeta^6, \zeta^7, \zeta^8, \zeta^9\right\}\right\}$$

4.2 The Chinese Remainder Theorem

Theorem 4.2.1 (Chinese Remainder Theorem)**.** If we are given $m, n \in \mathbb{Z}^+$ and $c, d \in \mathbb{Z}$. If $\gcd(m,n) = 1$, then there is a (actually, infinitely many) $z \in Z$ satisfying

$$z \equiv c \pmod{m} \text{ and } z \equiv d \pmod{n}$$

Proof. Since $\gcd(m,n) = 1$, there is a $\mathbb{Z}$-combination of m and n such that $am + bn = 1$ according to Euclid and theorem 2.4.1. Multiplying this $\mathbb{Z}$ combination by $c - d$, we have

$$(c-d)(am+bn) = (ca-da)m + (cb-db)n = c-d$$
$$(da-ca)m + (db-cb)n = d-c$$
$$(da-ca)m + c = (cb-db)n + d$$

Note that

$$(da-ca)m + c \equiv c \pmod{m} \quad \text{and} \quad (cb-db)n + d \equiv d \pmod{n}$$

and we've found a $z \in Z$ satisfying the requirement given by the question. In fact, any element from the set $\{z + kmn, k \in \mathbb{Z}\}$ satisfies the modulo relations, that's what I mean by the parenthesis in the question. □

Problem 4.2.1 (4.2.4)**.** Find all $z \in \mathbb{Z}$ such that

$$z \equiv 1 \pmod{5} \quad \text{and} \quad z \equiv 2 \pmod{7} \quad \text{and} \quad z \equiv 3 \mod 9$$

Solution 4.2.1. First we find a number a that satisfies $a \equiv 1 \pmod{5}$ and $7 \cdot 9 \mid a$. Similarly, we will then find b, a multiple of $5 \times 9 = 45$, that satisfies $y \equiv 2 \pmod{7}$. Finally, we will find c, a multiple of $5 \cdot 7 = 35$ that satisfies $c \equiv 3 \pmod{9}$. Note that a and $a + b + c$ have the same remainder when divided by 5 because $b + c$ is a multiple of 5. Likewise for b and c regarding remainders when divided by 7 and 9. After reducing three congruence modulo relations to two, the question becomes significantly easier. The smallest positive a, b, c are

$$a = 63 \cdot 2 = 126, b = 45 \cdot 3 = 135, \text{ and } c = 35 \cdot 6 = 210$$

Adding them together yields $126 + 135 + 210 = 471$. Since $471 > 5 \cdot 7 \cdot 9 = 315$, the smallest positive z that satisfies all three congruence modulo relation is $471 - 315 = 156$, and all $z \in \mathbb{Z}$ that satisfies these relation can be expressed as the elements of the set

$$\{156 + 315k : k \in \mathbb{Z}\}$$

4.3 Arithmetic in $\mathbb{Z}/n\mathbb{Z}$

Definition 4.3.1. For any $n \in \mathbb{Z}^+$ define $\mathbb{Z}/n\mathbb{Z}$ as

$$\mathbb{Z}/n\mathbb{Z} = \{[a]_n : a \in \mathbb{Z}\}$$

Definition 4.3.2. Fix an $n \in \mathbb{Z}^+$. Define addition and multiplication in $\mathbb{Z}/n\mathbb{Z}$ by

$$[a]_n + [b]_n = [a+b]_n \text{ and } [a]_n \cdot [b]_n = [ab]_n$$

Example 4.3.1. An extremely good application of $\mathbb{Z}/n\mathbb{Z}$ is to determine the remainders when divided by certain numbers. For example, we will find the remainder when 3624 is divided by 11:

$$\begin{aligned}[3624]_{11} &= [3\cdot 10^3 + 6\cdot 10^2 + 2\cdot 10 + 4]_{11} \\ &= [3]_{11}\cdot 10^3_{11} + [6]_{11}\cdot 10^2_{11} + [2]_{11}\cdot [10]^1_{11} + [4]_{11}\cdot [10]^0_{11} \\ &= [3]_{11}\cdot [-1]^3_{11} + [6]_{11}\cdot [-1]^2_{11} + [2]_{11}\cdot [-1]^1_{11} + [4]_{11}\cdot [-1]^0_{11} \\ &= -[3]_{11} + [6]_{11} - [2]_{11} + [4]_{11} \\ &= [-3+6-2+4]_{11} \\ &= [5]_{11}\end{aligned}$$

which shows exactly why the way to find the remainder when divided by 11 is by finding the difference between sums of alternate digits.

Proposition 4.3.3 (The Pigeonhole Principle)**.** Suppose we are given finite sets A and B with cardinality $|A|$ and $|B|$, respectively. Let $f : A \to B$ be a function.

(1) If f is injective then $|A| \leqslant |B|$.

(2) If f is surjective then $|A| \geqslant |B|$.

(3) If f is bijective then $|A| = |B|$.

(4) If $|A| = |B|$, then f is injective $\iff$ f is surjective.

Future reference: Chinese Remainder Theorem II, proposition 4.5.2, problem 5.1.2, theorem 4.6.11

Theorem 4.3.4 (Chinese Remainder Theorem II)**.** If $\gcd(m,n) = 1$, then the function

$$f : \mathbb{Z}/mn\mathbb{Z} \to \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$$

defined by $f([z]_{mn}) = ([z]_m, [z]_n)$ is a well-defined bijection.

Proof. Suppose $[x]_{mn} = [y]_{mn}$, then

$$x \equiv y \pmod{mn} \implies x \equiv y \pmod{m} \text{ and } x \equiv y \pmod{n}$$

i.e., $[x]_m = [y]_m$ and $[x]_n = [y]_n$, and we have shown that

$$f([x]_{mn}) = f([y]_m n) \implies ([x]_m, [x]_n) = ([y]_m, [y]_n)$$

Thus f is well-defined.

Knowing $|\mathbb{Z}/mn\mathbb{Z}| = mn = |\mathbb{Z}/m\mathbb{Z}|\cdot|\mathbb{Z}/n\mathbb{Z}|$, in order to prove f is a bijection, it suffices to show f is injective by proposition 4.3.3. Consider a, b such that $[a]_{mn} \neq [b]_{mn}$. Then either $a \not\equiv b \pmod{m}$ or $a \not\equiv b \pmod{n}$ (or both). In either case, $([a]_m, [a]_n) \neq ([b]_m, [b]_n)$. Therefore f is injective and thus bijective.

Future reference: proposition 4.6.5 □

Problem 4.3.1 (4.3.16, old version)**.** Call a function well-defined if $x = y \implies f(x) = f(y)$.

(1) Consider the function $f : \mathbb{Z}/3\mathbb{Z} \to \mathbb{Z}/9\mathbb{Z}$ by $f([a]_3) = [a^3]_9$. Show this function is well defined.

(2) Consider the function $f : \mathbb{Z}/3\mathbb{Z} \to \mu_9$ by $f([a]_3) = e^{2\pi i a^3/9}$. Show this function is defined.

Solution 4.3.1.

(1) If $[x]_3 = [y]_3$ then $x \equiv y \pmod 3$. Therefore there exists $k \in \mathbb{Z}$ such that $y = x + 3k$. Then,

$$\begin{aligned}\left[y^3\right]_9 = \left[(x+3k)^3\right]_9 &= \left[x^3 + 9x^2k + 27xk^2 + 27k^3\right]_9 \\ &= \left[x^3 + 9(x^2k + 3xk^2 + 3k^3)\right]_9 \\ &= \left[x^3\right]_9\end{aligned}$$

Therefore f is well-defined.

(2) Just like the previous part, suppose $[x]_3 = [y]_3$ then $\exists k \in \mathbb{Z}$ such that $y = x + 3k$. Then by (1), if we denote $x^k + 3xk^2 + 3k^3$ as n,

$$e^{2\pi i y^3/9} = e^{2\pi i(x^3+9n)/9} = e^{2\pi i x^3/9} \cdot e^{2n\pi i} = e^{2\pi i x^3/9}$$

which shows f is well-defined.

4.4 Rings and the Units of $\mathbb{Z}/n\mathbb{Z}$

Definition 4.4.1. A ***ring*** is a ordered triplet $(R, +, \cdot)$ where R is a set and

$$+ : R \times R \to R$$
$$\cdot : R \times R \to R$$

are functions with the following properties:

(1) $a + b = b + a$ for all $a, b \in R$

(2) $(a + b) + c = a + (b + c)$ for all $a, b, c \in R$

(3) There exists an ***additive identity***, 0_R, such that $a + 0_R = a$ for all $a \in R$

(4) For every $a \in R$, there exists an element $-a \in R$ such that $a + (-a) = 0_R$

(5) $(a \cdot b) \cdot c = a \cdot (b \cdot c)$ for all $a, b, c \in R$

(6) For all $a, b, c \in R$,

$$a \cdot (b + c) = a \cdot b + a \cdot c$$
$$(a + b) \cdot c = a \cdot c + b \cdot c$$

Proposition 4.4.2. A ring is closed under both addition and multiplication, i.e., for $a, b \in R$, we have $a + b \in R$ and $ab \in R$.

Usually, "let R be a ring" is a shorthand version of saying "let $(R, +, \cdot)$ be a ring. Both are correct. Different from addition and multiplication in $\mathbb{R}$, the following do NOT need to be satisfied in order for R to be a ring:

∗ $a \cdot b = b \cdot a$. Counterexample:

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \times \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \text{ but } \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \times \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

∗ The existence of a multiplication identity and multiplicative inverses. Counterexample: the set of even integers.

∗ Cancellation law: $a \cdot b = a \cdot c \implies b = c$. Counterexample:

$$[2]_6 \cdot [4]_6 = [2]_6 \cdot [1]_6 \text{ but } [4]_6 \neq [1]_6$$

Definition 4.4.3. A ***commutative ring*** R is a ring that satisfies $ab = ba$ for all $a, b \in R$.

Definition 4.4.4. A ***ring with 1*** is a ring with multiplicative identity: $\exists 1_R \in R$ such that $a \cdot 1_R = a = 1_R \cdot a$ for all $a \in R$.

Several propositions that might be helpful:

Proposition 4.4.5. For every $a \in R$ and the additive identity, 0_R, we have $a \cdot 0_R = 0 = 0 \cdot a_R$.

Proposition 4.4.6. A ring has a unique additive identity.

Proposition 4.4.7. Each $a \in R$ has a unique additive inverse.

Back to new concepts:

Definition 4.4.8. Let R be a ring with 1. An element $a \in R$ is called a ***unit*** if there exists a^{-1}, called the ***multiplicative inverse*** of a, such that $a \cdot a^{-1} = 1 = a^{-1} \cdot a$.

The set of all units in R is denoted as $R^{\times}$.

More propositions...

Proposition 4.4.9. If a set R has a multiplicative identity, it has a unique one.

Proposition 4.4.10. If $a \in R$ has a multiplicative inverse, then it has a unique one.

Future reference: problem 4.4.2

Proposition 4.4.11. Cancellation law applies if the canceled term is a unit. In other words, if $a, b \in R$ and $c \in R^{\times}$, then

$$ca = cb \iff a = b \text{ and } ac = bc \iff a = b$$

Future reference: Euler's Theorem

Example 4.4.1. Examples of $R^{\times}$: $\mathbb{R}^{\times} = \mathbb{R} \setminus \{0\}$; $\mathbb{C}^{\times} = \mathbb{C} \setminus \{0\}$;, $\mathbb{Z}^{\times} = \{-1, 1\}$; $Mn(\mathbb{R})^{\times} = \{A \in Mn(\mathbb{R}) : \det(A) \neq 0\}$, where $Mn(\mathbb{R})$ denotes the set of all $n \times n$ matrices with entries from $\mathbb{R}$.

Theorem 4.4.12. Fix any $n \in \mathbb{Z}$, the set of all units of the ring $\mathbb{Z}/n\mathbb{Z}$ is

$$(\mathbb{Z}/n\mathbb{Z})^{\times} = \{[a]_n : \gcd(a, n) = 1\}$$

Proof. First thing first, $[1]_n$ is always the unit for $\mathbb{Z}/n\mathbb{Z}$ by the definition of a multiplication identity. Now, if $[a]_n \in (\mathbb{Z}/n\mathbb{Z})^{\times}$, then there exists a $[b]_n \in \mathbb{Z}/n\mathbb{Z}$ such that $[a]_n \cdot [b]_n = [1]_n$.

Note that all elements of $[a]_n$ can be written in the form of $pn + a$ and all elements of $[b]_n$ can be written in the form of $qn + b$, where $p, q \in \mathbb{Z}$. Then,

$$\begin{aligned}
[a]_n \cdot [b]_n &= [ab]_n \\
&= [(pn + a)(qn + b)]_n \\
&= [pqn^2 + pbn + qan + ab]_n \\
&= [n(pqn + pb + qa)]_n + [ab]_n \\
&= [ab]_n
\end{aligned}$$

If $[ab]_n = [1]_n$, then $n \mid ab - 1$ which immediately implies $\gcd(a, n) \mid ab - 1$. However, $\gcd(a, n)$ clearly divides a and ab as well. To both divide $ab - 1$ and ab, we conclude that $\gcd(a, b) = 1$, which finishes the proof.

Future reference: problem 4.4.1, Fermat's Little Theorem, problem 4.5.1, Euler's Theorem □

Corollary 4.4.12.1. If p is prime then

$$(\mathbb{Z}/p\mathbb{Z})^{\times} = \{[1]_p, [2]_p, \cdots, [p - 1]_p\}$$

Future reference: problem 4.4.2

Problem 4.4.1 (4.4.23, old version)**.** Suppose p is a prime and $n \in \mathbb{Z}^+$. Find a formula for $|(\mathbb{Z}/p^n\mathbb{Z})^\times|$.

Solution 4.4.1. By theorem 4.4.12, all elements $a \in (\mathbb{Z}/p^n\mathbb{Z})$ satisfy $\gcd(a, p^n) = 1$. By problem 3.2.1, we know that there are $p^n - p^{n-1}$ such a's. Therefore $|(\mathbb{Z}/p^n\mathbb{Z})^\times| = p^n - p^{n-1}$.

Future reference: proposition 4.6.10

Problem 4.4.2 (4.4.28, old version)**.** Prove Wilson's Theorem: p is a prime if and only if

$$(p-1)! \equiv -1 \pmod p$$

Solution 4.4.2. When $p = 2$, this statement is obviously true since $1! = 1$ and $1 \equiv -1 \pmod 2$. Now, besides 2, all prime numbers are odd, thus all $p - 1$ are even. Note that

$$\begin{aligned}\prod_{i=1}^{p-1}[i]_n = [1]_n \cdot [p-1]_n \cdot \prod_{i=2}^{p-2}[i]_n &= [p-1]_n \cdot \prod_{i=2}^{p-2}[i]_n \\ &= [-1]_n \cdot \prod_{i=2}^{p-2}[i]_n\end{aligned}$$

By corollary 4.4.12.1, since P is a prime, $[i]_n \in (\mathbb{Z}/n\mathbb{Z})^\times$ for all i between 1 and $p - 1$. Then, by proposition 4.4.10, each $[j]_n$ has a unique multiplicative inverse $[k]_n$. Moreover, $[j]_n$ itself is the unique multiplicative inverse of $[k]_n$. Therefore, since $(\mathbb{Z}/p\mathbb{Z})^\times \setminus \{[1]_n, [p-1]_n\}$ has $p - 3$ elements, these elements will form $\dfrac{p-3}{2}$ pairs, each of which has a product of $[1]_n$. Their product is, of course, still $[1]_n$. Since

$$\prod_{i=1}^{p-1}[i]_n = [-1]_n \cdot [1]_n = [-1]_n$$

we have proven Wilson's Theorem. □

4.5 Fermat's Little Theorem

Proposition 4.5.1. Let R be a ring with 1. If $a, b \in R^{\times}$ then $ab \in R^{\times}$.

Proof. We will prove a stronger statement: that $ab \in R^{\times}$ and $ab^{-1} = b^{-1}a^{-1}$. Two lines are enough.

$$(ab)(b^{-1}a^{-1}) = abb^{-1}a^{-1} = a(bb^{-1})a^{-1} = a(1_R)a^{-1} = aa^{-1} = 1_R$$
$$(b^{-1}a^{-1})(ab) = b^{-1}a^{-1}ab = b^{-1}(a^{-1}a)b = b^{-1}(1_R)b = b^{-1}b = 1_R$$

Future reference: proposition 4.5.2, Fermat's Little Theorem, Euler's Theorem □

Proposition 4.5.2. Let R be a ring $\mathbb{Z}/n\mathbb{Z}$ where $n \in Z^{+}$. Define a function

$$\text{mult}_u : R^{\times} \to R^{\times}$$

by $\text{mult}_u([x]_n) = [u]_n \cdot [x]_n$. This is well defined by proposition 4.5.1. For any $u \in R^{\times}$, mult_u is a bijection.

Proof. Since the domain and the codomain are the same, they have equal cardinality. Part (4) of the The Pigeonhole Principle suggests that to prove mult_u is a bijection, it suffices to show it is an injection, i.e., for $0 \leqslant x_1, x_2 \leqslant n-1$,

$$[x_1]_n \neq [x_2]_n \implies \text{mult}_u([x_1]_n) \neq \text{mult}_u([x_2]_n)$$

Since $u \in R^{\times}$, we know $\gcd(u, n) = 1$. Note that

$$\text{mult}_u([x_1]_n) = [x_1]_n \cdot [u]_n = [ux_1]_n \text{ and}$$
$$\text{mult}_u([x_2]_n) = [x_2]_n \cdot [u]_n = [ux_2]_n$$

If $[x_1]_n \neq [x_2]_n$, then $x_1 \neq x_2$ and $ux_1 \neq ux_2$. If we assume $[ux_1]_n = [ux_2]_n$ then $n \mid (ux_1 - ux_2)$. Since $\gcd(n, u) = 1$, it must be the case that $n \mid (x_1 - x_2)$ which is a contradiction, given that $x_1 \neq x_2$ and $0 \leqslant x_1, x_2 \leqslant n-1$. Therefore the assumption that $x_1 \neq x_2$ and $\text{mult}_u(x_1) = \text{mult}_u(x_2)$ is false, and the function is indeed injective and bijective.

Future reference: Fermat's Little Theorem, Euler's Theorem □

Theorem 4.5.3 (Fermat's Little Theorem)**.** Suppose p is a prime and $\gcd(a, p) = 1$ (or $p \nmid a$), then

$$a^{p-1} \equiv 1 \pmod{p}$$

Future reference: problem 4.5.1

Proof. Since $\gcd(a, p) = 1$, by theorem 4.4.12, $[a] \in (\mathbb{Z}/p\mathbb{Z})^{\times}$. Define a function

$$\text{mult}_{[a]} : (\mathbb{Z}/p\mathbb{Z})^{\times} \to (\mathbb{Z}/p\mathbb{Z})^{\times} \text{ by } \text{mult}_{[a]}([x]) = [a][x].$$

By proposition 4.5.2, $\text{mult}_{[a]}$ is a bijection. Therefore, the multiplication by $[a]$ permutes the elements of $(\mathbb{Z}/p\mathbb{Z})^{\times}$. In other words each element of $(\mathbb{Z}/p\mathbb{Z})^{\times}$ appears exactly once as image and exactly once as preimage as well. Then,

$$\{[1], [2], \cdots, [p-1]\} = \{[1a], [2a], [3a]\cdots, [(p-1)a]\}$$

It follows that
$$\prod_{i=1}^{p-1}[i] = \prod_{i=1}^{p-1}[ai] \implies [(p-1)!] = [a^{(p-1)}]\cdot[(p-1)!]$$
Proposition 4.5.1 shows that the products of units are units. Since p is a prime, $[(p-1)!]$ is therefore a product of units for $\mathbb{Z}/p\mathbb{Z}$. Thus $[(p-1)!]$ has a multiplicative inverse. Multiplying both sides of the equation with this inverse, we have

$$\begin{aligned}[(p-1)!]\cdot[(p-1)!]^{-1} &= [a^{(p-1)}]\cdot[(p-1)!]\cdot[(p-1)!]^{-1}\\ [1] &= [a^{p-1}]\end{aligned}$$

and we've proven Fermat's Little Theorem.

Future reference: Euler's Theorem □

Problem 4.5.1 (4.5.6)**.** Let p be a prime.

(1) Suppose $0 < k < p$. Prove $\binom{p}{k} \equiv 0 \pmod{p}$.

(2) Deduce $(x+y)^p \equiv x^p + y^p \pmod{p}$ for all $x, y \in \mathbb{Z}$

(3) Use (2) to prove $n^p \equiv n \pmod{p}$ for all $n \in \mathbb{Z}^+$.

(4) Use (3) to give a new proof of the Fermat's Little Theorem.

Solution 4.5.1.

(1) Rewrite $\binom{p}{k}$ as
$$\binom{p}{k} = \frac{p!}{k!(p-k)!}$$
Clearly p divides the numerator. However, since p is prime, $k \nmid p$ for all $k \in (0,p)$. Therefore p does not divide the numerator, and $\binom{p}{k}$ is a multiple of p, i.e., $\binom{p}{k} \equiv 0 \pmod{p}$.

(2) Apply binomial theorem to $(x+y)^p$:
$$(x+y)^p = \sum_{i=0}^{p}\binom{p}{i}x^{p-i}y^i = x^p + \sum_{i=1}^{p}\binom{p}{i}x^{p-1}y^i + y^p$$
Since $\binom{p}{k} \equiv 0 \pmod{p}$ for all $i \in (0,k)$ by (1), p divides the second term. Thus
$$(x+y)^p \equiv x^p + y^p \pmod{p}$$

(3) Since we are trying to prove $n^p \equiv n \pmod{p}$ for all $n \in \mathbb{Z}^+$, the base case would be $n = 1$:
$$1^p = 1 \text{ so } 1^p \equiv 1 \pmod{p}.$$
Now for the inductive step, assume $k^p \equiv k \pmod{p}$ for some $k \in \mathbb{Z}^+$. By (2), we have
$$(k+1)^p \equiv k^p + 1^p \equiv k^p + 1 \pmod{p}$$
which finishes the proof of the inductive step.

Therefore, $n^p \equiv n$ for all $n \in \mathbb{Z}^+$.

(4) For convenience purposes, we will denote $[x]_p$ as $[x]$. Then we can rewrite the statement of (3) as

$$[n]^p = [n]$$

Fermat's Little Theorem states that for all $n \in \mathbb{Z}$, if $\gcd(n, p) = 1$ then $[n]^{p-1} = [1]$. By theorem 4.4.12, we have $n \in \mathbb{Z}/p\mathbb{Z}$. Therefore there exists a multiplicative inverse of $[n]$ which we will denote as $[n]^{-1}$. Multiplying both sides of the equation above by $[n]^{-1}$ yields

$$\begin{aligned}[n]^p \cdot [n]^{-1} &= [n] \cdot [n]^{-1} \\ [n]^{p-1} \cdot [n] \cdot [n]^{-1} &= [1] \\ [n]^{p-1} &= [1]\end{aligned}$$

since $[n] \cdot [n]^{-1} = 1$. Therefore we've proven Fermat's Little Theorem using a different method.

4.6 Isomorphism and Euler's Totient Function

Definition 4.6.1. Suppose S and T are rings. The ***product ring*** $S \times T$ is the ring whose elements are ordered pairs (s,t) with $s \in S$ and $t \in T$. Addition and multiplication in $S \times T$ are defined by

$$(s_1,t_1) + (s_2,t_2) = (s_1 + s_2, t_1 + t_2)$$
$$(s_1,t_1) \cdot (s_2,t_2) = (s_1 \cdot s_2, t_1 \cdot t_2)$$

Proposition 4.6.2. If S and T are rings with 1, then the product ring $S \times T$ is again a ring with 1.

Proof. If S and T are both rings with 1 then for any $s \in T$ and $t \in T$ we have

$$(s,t) \cdot (1_S, 1_T) = (s \cdot 1_S, t \cdot 1_T) = (s,t).$$

Therefore $(1_S, 1_T) = 1_{S\times T}$. Hence proven.

Future reference: proposition 4.6.8 □

Definition 4.6.3. Suppose R and S are rings. A function $f : R \to S$ is a ***ring homomorphism*** if for all $a, b \in R$,

$$f(a+b) = f(a) + f(b)$$
$$f(a \cdot b) = f(a) \cdot f(b)$$

Definition 4.6.4. A ***ring isomorphism*** is a bijective ring homomorphism.

Example 4.6.1. Consider the function $f : \mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$ defined by $f(x) = [x]_n$. Clearly, for $a, b \in \mathbb{Z}$,

$$f(a+b) = [a+b]_n = [a]_n + [b]_n = f(a) + f(b)$$
$$f(a \cdot b) = [a \cdot b]_n = [a]_n \cdot [b]_n = f(a) \cdot f(b)$$

But f is not bijective since $f(0) = f(n)$. Thus f is a ring homomorphism but not ring isomorphism.

Proposition 4.6.5. Suppose $a, b \in \mathbb{Z}$ and $\gcd(a,b) = 1$. Then the function

$$f : \mathbb{Z}/ab\mathbb{Z} \to \mathbb{Z}/a\mathbb{Z} \times \mathbb{Z}/b\mathbb{Z}$$

defined by $f([x]_{ab}) = ([x]_a, [x]_b)$ is a ring isomorphism.

Proof. We first show that f is a ring homomorphism:

$$f([x]_{ab} + [y]_{ab}) = f([x+y]_{ab}) = ([x+y]_a, [x+y]_b) = ([x]_a, [x]_b) + ([y]_a, [y]_b) = f([x]_{ab}) + f([y]_{ab})$$
$$f([x]_{ab} \cdot [y]_{ab}) = f([x \cdot y]_{ab}) = ([x \cdot y]_a, [x \cdot y]_b) = ([x]_a, [x]_b) \cdot ([y]_a, [y]_b) = f([x]_{ab}) \cdot f([y]_{ab})$$

Then, by Chinese Remainder Theorem II, we also know that f is a bijection. Hence f is a ring isomorphism.

Future reference: theorem 4.6.11 □

Proposition 4.6.6. Suppose R and S are rings with 1, and $f : R \to S$ is a ring isomorphism satisfying $f(1_R) = 1_S$. Then for every $r \in R$,

$$r \in R^\times \iff f(r) \in S^\times$$

Proof. To show $\implies$, suppose $r \in R^\times$, then there exist a r^{-1} such that $rr^{-1} = 1_R$. Since f is a ring isomorphism, it follows that $f(1_R) = 1_S \implies f(r) \cdot f(r^{-1}) = 1_S$. Therefore $f(r) \in S$ has a multiplicative inverse, i.e., $f(r) \in S^\times$.

To show $\impliedby$, suppose $f(r) \in S^\times$. Then it has a multiplicative inverse which we call $(f(r))^{-1} \in S^\times$. Since f is a ring isomorphism, f is bijective and thus surjective. Therefore there exists an element $r' \in R$ such that $f(r') = (f(r))^{-1}$. Then, by the definition of ring homomorphism,

$$1_S = f(r) \cdot (f(r))^{-1} = f(r) \cdot f(r') = f(rr')$$

which implies $rr' = 1_R$. Therefore $r \in R^\times$. □

Proposition 4.6.7. Suppose R and S are rings with 1, and $f : R \to S$ is a ring isomorphism satisfying $f(1_R) = 1_S$. Then f restricted to $R^\times$ is a bijection $f : R^\times \to S^\times$.

Future reference: theorem 4.6.11

Proposition 4.6.8. Suppose S and T are rings with 1. Then $(S \times T)^\times = S^\times \times T^\times$. In other words, for $s \in S$ and $t \in T$, (s,t) is a unit in $S \times T$ if and only if $s \in S^\times$ and $T \in T^\times$.

Proof. We will first show $\implies$, i.e., $(s,t) \in (S \times T)^\times \implies s \in S^\times$ and $t \in T^\times$. Suppose $(s,t) \in (S \times T)^\times$, then there exist $s' \in S$ and $t' \in T$ such that $(s,t) \cdot (s',t') = 1_{S\times T}$. As shown in proposition 4.6.2, $1_{S\times T} = (1_S, 1_T)$. Therefore $ss' = 1_S \implies s \in S^\times$ and $tt' = 1_T \implies t \in T^\times$.

The $\impliedby$ direction is just obvious. If $s \in S^\times$ and $t \in T^\times$ then there exist $s^{-1} \in S$ and $t^{-1} \in T$ such that $ss^{-1} = 1_S$ and $tt^{-1} = 1_T$. Then

$$(s,t) \cdot (s^{-1}, t^{-1}) = (ss^{-1}, tt^{-1}) = (1_S, 1_T) = 1_{S\times T}$$

which suggests $(s,t) \in (S \times T)^\times$.

Future reference: theorem 4.6.11 □

Having introduced the basics of ring homomorphism and isomorphism, now we can move to Euler's totient function, also known as Euler's φ function, which addresses a simple question: given a positive integer n, how many positive integers less than n are co-prime with n?

Definition 4.6.9 (Euler's totient function)**.** Define Euler's totient function $\varphi : \mathbb{Z}^+ \to \mathbb{Z}^+$ by $\varphi(n) = |(\mathbb{Z}/n\mathbb{Z})^\times|$.

Proposition 4.6.10. By problem 4.4.1, if p is a prime and $n \in Z^+$, then $|(\mathbb{Z}/n\mathbb{Z})^\times| = p^n - p^{n-1}$.

Theorem 4.6.11. Suppose $a, b \in \mathbb{Z}^+$ are co-prime. Then $\varphi(ab) = \varphi(a) \cdot \varphi(b)$.

Proof. The function $f : \mathbb{Z}/ab\mathbb{Z} \to \mathbb{Z}/a\mathbb{Z} \times \mathbb{Z}/b\mathbb{Z}$ defined by $f([x]_{ab}) = ([x]_a, [x]_b)$ is a ring isomorphism by proposition 4.6.5. In addition, since $f([1]_{ab}) = ([1]_a, [1]_b)$, by proposition 4.6.7 we know f restricted to $(\mathbb{Z}/ab\mathbb{Z})^\times$ is a bijection $f : (\mathbb{Z}/ab\mathbb{Z})^\times \to (\mathbb{Z}/a\mathbb{Z} \times \mathbb{Z}/b\mathbb{Z})^\times$. Therefore,

$$\begin{aligned}
\varphi(ab) &= |(\mathbb{Z}/ab\mathbb{Z})^\times| \\
&= |(\mathbb{Z}/a\mathbb{Z} \times \mathbb{Z}/b\mathbb{Z})^\times| && \text{(bijection $\iff$ same cardinality, proposition 4.3.3)} \\
&= |(\mathbb{Z}/a\mathbb{Z})^\times \times (\mathbb{Z}/b\mathbb{Z})^\times| && \text{(Proposition 4.6.8)} \\
&= |(\mathbb{Z}/a\mathbb{Z})^\times| \cdot |(\mathbb{Z}/b\mathbb{Z})^\times| && \text{(This holds because both sets are finite)} \\
&= \varphi(a) \cdot \varphi(b)
\end{aligned}$$

which completes the proof.

Future reference: theorem 4.6.1 □

Problem 4.6.1 (4.6.11)**.** Suppose $n \in \mathbb{Z}^+$ has prime factorization $n = \prod_{i=1}^{s} p_i^{e_i} = p_1^{e_1} p_2^{e_2} \cdots p_s^{e_s}$. Show that

$$\varphi(n) = n \cdot \prod_{i=1}^{s} \left(1 - \frac{1}{p_i}\right)$$

Solution 4.6.1. Notice that after being prime factorized, n is now expressed as the product of s pairwise co-prime positive integers, each equaling to a prime raised to some positive power. Therefore,

$$\begin{aligned}
\varphi(n) &= \prod_{i=1}^{s} \varphi\left(p_i^{e_i}\right) = \varphi\left(p_1^{e_1}\right) \cdot \varphi\left(p_2^{e_2}\right) \cdots \varphi\left(p_s^{e_s}\right) \\
&= \prod_{i=1}^{s} \left(p_i^{e_i} - p_i^{e_i - 1}\right) = \left(p_1^{e_1} - p_1^{e_1 - 1}\right)\left(p_2^{e_2} - p_2^{e_2 - 1}\right) \cdots \left(p_s^{e_s} - p_s^{e_s - 1}\right) && \text{(by proposition 4.6.10)} \\
&= \prod_{i=1}^{s} \left(p_i^{e_i}\left(1 - \frac{1}{p_i}\right)\right) = \left(p_1^{e_1}\left(1 - \frac{1}{p_1}\right)\right)\left(p_2^{e_1 2}\left(1 - \frac{1}{p_2}\right)\right) \cdots \left(p_s^{e_s}\left(1 - \frac{1}{p_s}\right)\right) \\
&= \left(\prod_{i=1}^{s} p_i^{e_i}\right) \cdot \prod_{i=1}^{s} \left(1 - \frac{1}{p_i}\right) = n \cdot \prod_{i=1}^{s} \left(1 - \frac{1}{p_i}\right)
\end{aligned}$$

Hence proven. □

4.7 Euler's Theorem

Recall that proposition 4.5.2 states that, given a congruence ring $\mathbb{Z}/n\mathbb{Z}$ where $n \in \mathbb{Z}^+$, the function

$$\text{mult}_u : (\mathbb{Z}/n\mathbb{Z})^\times \to (\mathbb{Z}/n\mathbb{Z})^\times$$

defined by $\text{mult}_u([x]_n) = [u]_n \cdot [x]_n$ is a bijection. Euler extended Fermat's Little Theorem a little bit, but the main idea are exactly the same: the bijection mult_u and the permutation of elements in $(\mathbb{Z}/n\mathbb{Z})^\times$ that can be used for cancellation.

Theorem 4.7.1 (Euler's Theorem)**.** (Euler, 1736) Suppose $a \in \mathbb{Z}$ and $n \in \mathbb{Z}^+$ such that $\gcd(a, n) = 1$. Then

$$a^{\varphi(n)} \equiv 1 \pmod n$$

Proof. Very similar to the proof of Fermat's Little Theorem.

Since $\gcd(a, n) = 1$, we know $[a] \in (\mathbb{Z}/n\mathbb{Z})^\times$ by theorem 4.4.12. Define a function

$$\text{mult}_{[a]} : (\mathbb{Z}/n\mathbb{Z})^\times \to (\mathbb{Z}/n\mathbb{Z})^\times \text{ by } \text{mult}_{[a]}([x]) = [a][x].$$

By proposition 4.5.2, $\text{mult}_{[a]}$ is a bijection. Therefore, the multiplication by $[a]$ permutes the elements of $(\mathbb{Z}/n\mathbb{Z})^\times$. In other words each element of $(\mathbb{Z}/n\mathbb{Z})^\times$ appears exactly once as image and exactly once as preimage as well, i.e., if we denote the set of all elements in $(\mathbb{Z}/n\mathbb{Z})^\times$ as

$$(\mathbb{Z}/n\mathbb{Z})^\times = \left\{[u_1], [u_2], \cdots, [u_{\varphi(n)}]\right\},$$

then

$$\left\{[a]\cdot[u_1], [a]\cdot[u_2], \cdots, [a]\cdot[u_{\varphi(n)}]\right\} = \left\{[u_1], [u_2], \cdots, [u_{\varphi(n)}]\right\}.$$

Multiplying all of the elements in the first set together, multiplying all of the elements in the second set together, and then setting the results equal to one another, we have

$$[a]^{\varphi(n)} \prod_{i=1}^{\varphi(n)} [u_i] = \prod_{i=1}^{\varphi(n)} [u_i].$$

Proposition 4.5.1 claims that products of units are still units. Therefore $\prod_{i=1}^{\varphi(n)} [u_i] \in (\mathbb{Z}/n\mathbb{Z})^\times$ and cancellation law for units holds, and

$$[a]^{\varphi(n)} = [1], \text{ i.e., } a^{\varphi(n)} \equiv 1 \pmod n.$$

□

Remark. If p is a prime then $\varphi(p) = p - 1$, which is exactly what Fermat's Little Theorem is about.

5 Polynomial Arithmetic

5.1 Integral Domains and Fields

Definition 5.1.1. The ***trivial ring*** is the ring $\{0\}$ consisting of a single element, 0, with addition and multiplication defined by $0 + 0 = 0$ and $0 \cdot 0 = 0$.

- $\ast$ The trivial ring is a ring with 1 since $0 \cdot a = a = a \cdot 0$ for every $a \in \{0\}$.
- $\ast$ Let R be a ring with 1. If $1_R = 0_R$, then R is the trivial ring.

Definition 5.1.2. An ***integral domain*** R is a nontrivial commutative ring with 1, such that for all $a, b \in R$,

$$ab = 0_R \implies a = 0_R \vee b = 0_R$$

Contrapositive:

$$a \neq 0_R \wedge b \neq 0_R \implies ab \neq 0_R$$

Definition 5.1.3. (Not in lecture notes) In a ring R, a non-zero element $a \in R$ is called a ***left zero divisor*** if there exists another non-zero element $b \in R$ such that $ab = 0_R$. Likewise for the definition of a ***right zero divisor***. An integral domain is a ring without zero divisors.

Definition 5.1.4. (Not in lecture notes) Let R be a ring. A ***nilpotent*** element of R is an element $x \in R$ such that there exists an $n \in \mathbb{N}$ and $x^n = 0$.

Remark. Being nilpotent is different from being a zero divisor.

If $x \in R$ is nilpotent then x is a zero divisor. This is because $x^n = 0 \implies x \cdot x^{n-1} = 0$.

However, a zero divisor might not be nilpotent. For example, in $\mathbb{Z}/6\mathbb{Z}, 2$ and 3 are both zero divisors but neither are nilpotent.

Example 5.1.1.

- $\ast$ The rings $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$, and $mathbbC$ are integral domains.
- $\ast$ The ring $\mathbb{Z}/12\mathbb{Z}$ isn't. For example, $[3]_{12} \cdot [4]_{12} = [0]_{12}$, in which case $[3]_{12}$ and $[4]_{12}$ are both zero divisors of $\mathbb{Z}/12\mathbb{Z}$.
- $\ast$ The ring $M_2(\mathbb{R})$ isn't. First, it is not commutative. Second,

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Theorem 5.1.5 (Cancellation law)**.** Let R be an integral domain and suppose $ab = ac$ for some $a, b, c \in R$. If $a \neq 0_R$ then $b = c$.

Proposition 5.1.6. A ring $R \neq \{0\}$ is an integral domain if and only if the cancellation law holds in R.

Proof.

$\Longrightarrow$: for $a, b, c \in R$, if $ab = ac$ and $a \neq 0$, then $ab - ac = a(b-c) = 0$. Since $a \neq 0$ and R is an integral domain, $b - c = 0$. Thus $b = c$.

$\Longleftarrow$: for $a, b \in R$, if $ab = 0$ but $a \neq 0$, then by cancellation law $ab = a \cdot 0 \implies b = 0$. Therefore at least one element between the two is 0. Thus R is an integral domain. □

Definition 5.1.7. A ***field*** is a nontrivial commutative ring with 1 in which every nonzero element has a multiplicative inverse.

In other words, if F is a nontrivial commutative ring with 1 then F is a field if and only if $F^\times = F \setminus \{0\}$.

Proposition 5.1.8. If F is a field then F is an integral domain.

Proof. We will denote 0_F as 0 for simplicity. Suppose $x, y \in F$ and $xy = 0$. If $x = 0$ or $y = 0$ (or both) then we are immediately done.

Suppose $x \neq 0$ and $y \neq 0$, there must exist an x^{-1} since F is a field. Then,

$$xy = 0 \implies x^{-1}xy = x^{-1}0 \implies y = 0$$

which contradicts $y \neq 0$. Therefore it is impossible for both x and y to be nonzero. Thus F is an integral domain.

Future reference: theorem 5.1.16 □

Proposition 5.1.9. $\mathbb{Z}/p\mathbb{Z}$ is a field (and, of course, integral domain) if and only if p is a prime since this is the only occasion in which $\mathbb{Z}/p\mathbb{Z}$ has no zero divisors.

Future reference: Gauss's Lemma I

Definition 5.1.10. Let R be a ring. We will denote the ring of polynomials with coefficients in R as $R[x]$. Similarly, $\mathbb{R}[x], \mathbb{Q}[x], \mathbb{Z}[x]$, and $\mathbb{C}[x]$ are the rings with real, rational, integer, and complex coefficients, respectively. (Of course there are more!)

Definition 5.1.11. If F is a field, then we denote by $F(x)$ the field of ***rational functions*** with coefficients in F:

$$F(x) = \left\{ \frac{a(x)}{b(x)} : a, (x), b(x) \in F[x] \text{ and } b(x) \neq 0 \right\}$$

Proposition 5.1.12. Suppose

$$f(x) = \sum_{i=0}^{n} a_i x^i = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$$

then we call n the ***degree*** of f, denoted as $\deg(x)$. By convention, the zero polynomial has degree $-\infty$. Therefore the set of all degrees of $f(x) \in R[x]$ is $\{-\infty\} \cup \mathbb{Z}^{\geqslant 0}$.

Future reference: Division Algorithm for polynomials

Definition 5.1.13. Let R be a commutative ring with 1 and suppose $a(x), b(x) \in R[x]$. We say $a(x)$ divides $b(x)$, and write $a(x) \mid b(x)$, if there is a $q(x) \in R[x]$ such that $b(x) = a(x) \cdot q(x)$.

Proposition 5.1.14. Suppose $f(x), g(x) \in R[x]$. If all coefficients of f and g are in an integral domain, then

$$\deg(f \cdot g) = \deg(f) + \deg(g)$$

Future reference: problem 5.1.1, Division Algorithm for polynomials, proposition 5.4.4, Eisenstein's Criterion

Remark. The converse is not always true. For example, $(\mathbb{Z}/12\mathbb{Z})[x]$ is not an integral domain, but it can sometimes still satisfy $\deg(f \cdot g) = \deg(f) \cdot \deg(g)$ where $f(x), g(x) \in \mathbb{Z}/12\mathbb{Z}$. If we let

$$f(x) = [3]x^5 + [1] \text{ and } g(x) = [5]x^2 + [1],$$

then

$$\begin{aligned} ([3]x^5 + [1])([5]x^2 + [1]) &= [15]x^7 + [3]x^5 + [5]x^2 + [1] \\ &= [3]x^7 + [3]x^5 + [5]x^2 + [1] \\ \deg(f \cdot g) = 7, \deg(f) = 5,& \text{ and } \deg(g) = 2 \end{aligned}$$

Of course, sometimes the degree equation can be false. For example, if we change change $g(x)$ to $([4]x^2 + [1])$, then

$$\begin{aligned} ([3]x^5 + [1])([4]x^2 + [1]) &= [12]x^7 + [3]x^5 + [4]x^2 + [1] \\ &= [3]x^5 + [4]x^2 + [1] \end{aligned}$$

which results in a 5^{th} degree function because the product of the coefficients of f's and g's leading terms equals $[0]$. If this product does not equal $[0]$, then $\deg(f \cdot g) = \deg(f) + \deg(g)$ holds even when $f(x), g(x) \in R[x]$ that's not an integral domain.

Proposition 5.1.15. If R is an integral domain and $a(x), b(x) \in R[x]$ with $b(x) \neq 0$, then

$$a(x) \mid b(x) \implies \deg(a) \leqslant \deg(b)$$

Remark. $b(x) \neq 0$ is important here because, by definition, if $b(x) = p(x) = 0$ then
$a(x) \cdot p(x) = b(x) \implies a(x) \mid b(x)$, but $a(x)$ can be of any degree.

Theorem 5.1.16. If R is an integral domain, then so is $R[x]$.

Proof. To prove $R[x]$ is an integral domain, we want to show that for $f(x), g(x) \in R$, if
$f(x) \cdot g(x) = 0$, then they cannot both be nonzero.

First, $f(x)$ and $g(x)$ has the form

$$\begin{aligned} f(x) &= \sum_{i=0}^{m} a_i x^i = a_m x^m + a_{m-1}x^{m-1} + \cdots + a_1 x + a_0 \\ g(x) &= \sum_{i=0}^{n} b_i x^i = b_n x^n + b_{n-1}x^{n-1} + \cdots + b_1 x + b_0 \end{aligned}$$

and the only case in which such polynomials equal 0 is when all coefficients are 0.

Suppose both $f(x)$ and $g(x)$ are nonzero, then both have nonzero term with lowest degree. (For example, the nonzero term with lowest degree of the polynomial $x^7 + 5x^6 + 8x^4$ is $8x^4$.) Suppose such term of $f(x)$ is $a_j x^j$ and such term of $g(x)$ is $b_k x^k$. Now look at the term x^{j+k} in the polynomial $f(x) \cdot g(x)$. The coefficient of this term is the sum of a bunch of ab's such that the index number of a plus the index number of b equals $j+k$. Among all these ab's, only $a_j b_k \neq 0$ because for every other ab, either the index number of a is smaller than j or that of b is smaller than k, but all of these a's and b's are zero because we've said $a_j x^j$ and $b_k x^k$ are the two nonzero terms with lowest degree. Therefore the term x^{j+k} has coefficient $a_j b_k$, which is nonzero by the contrapositive of the defnition of integral domain. Now that we've have a nonzero coefficient in $f(x) \cdot g(x)$, we are done with the proof. If $f(x) \neq 0$ and $g(x) \neq 0$, then $f(x) \cdot g(x) \neq 0$. $R[x]$ is an integral domain. Hence proven.

Future reference: problem 5.1.1, Fundamental Theorem of Arithmetic for Polynomials, Gauss's Lemma I □

Problem 5.1.1 (5.1.16)**.** Prove that every finite integral domain is a field.

Solution 5.1.1–1. From the hint provided in the text: let R be an integral domain. Choose a nonzero element $r \in R$ and consider the function $\text{mult}_r : R \to R$. First of all, since the function maps R to R, the cardinality of domain equals that of the codomain, i.e., $|R| = |R|$. Since we are given that R is a finite integral domain, by the The Pigeonhole Principle, showing mult_r is injective implies mult_r is bijective.

Now suppose $f(x) = f(y)$ for some $x, y \in R$. Then

$$rx = ry \implies rx - ry = 0 \implies r(x-y) = 0$$

We've set r to be nonzero, but R is an integral domain, so $(x-y)$ must equal 0, i.e., $x = y$. Therefore $f(x) = f(y) \implies x = y$ and f is an injection and, more generally, a bijection.

Also, since R is an integral domain and f is a bijection, there exists some s such that $f(s) = rs = 1$. Therefore r has a multiplicative inverse. Since we've chosen r randomly, we've shown every nonzero $r \in R$ has a multiplicative inverse, i.e., R is a field. Hence proven. □

Solution 5.1.1–2. Alternative solution: again, consider any nonzero element $r \in R$. Then the set

$$\left\{r^k : k \in \mathbb{Z}^{\geqslant 0}\right\} = \left\{1, r, r^2, r^3, \cdots\right\}$$

must contain a finite amount of distinct elements since R is finite. Therefore there exists m and n such that $0 < m < n$ and $r^m = r^n$, and we have $r^m \cdot 1 = r^m = r^n = r^m \cdot r^{n-m}$. (Of course, we denote 1_R as 1.) In an integral domain, since r is nonzero, we claim r^2 is nonzero, and if we assume r^j is nonzero then r^{j+1}, a product of nonzero elements r and r^j, is also nonzero. Therefore all element in the set above are nonzero by induction. By Cancellation law we have

$$r^m \cdot r^{n-m} = r^m \cdot 1 \implies r^{n-m} = 1$$

Finally, since $r^{n-m} = r \cdot r^{n-m-1}$, we've found a multiplicative inverse of r. Therefore for any nonzero $r \in R$, there exists a multiplicative inverse, and R is a field. □

Problem 5.1.2 (5.1.17)**.** Show that if R is an integral domain then $R^\times = R[x]^\times$, i.e., the units in $R[x]$ are precisely the units in R, viewed as constant polynomials.

Solution 5.1.2. To show $R^\times = R[x]^\times$, it suffices to show both $R^\times \subseteq R[x]^\times$ and $R[x]^\times \subseteq R^\times$.

The first one is obvious since for all $a \in R^\times$, $f(x) = a$ is a zero-degree polynomial as well as a unit of $R[x]$.

Now look at the second statement. Suppose $f(x), g(x) \in R[x]^\times$, then $f(x) \cdot g(x) = 1$, which has degree 0. By theorem 5.1.16, we know R is an integral domain implies $R[x]$ also is. By proposition 5.1.14 we have

$$\deg(f \cdot g) = 0 = \deg(f) + \deg(g) \implies \deg(f) = \deg(g) = 0$$

Then both $f(x)$ and $g(x)$ are constants, and their product is 1. Therefore these constants are units of R, and we've also shown $R[x]^\times \subseteq R[x]$. □

5.2 The Division Algorithm for Polynomials

Theorem 5.2.1 (Division Algorithm for polynomials)**.** Suppose F is a field and $a(x), b(x) \in F[x]$ with $b(x) \neq 0$. Then there are unique $q(x), r(x) \in F[x]$ such that

$$a(x) = b(x)q(x) + r(x) \text{ and } \deg(r) < \deg(b)$$

Proof. The proof consists of two parts: showing $q(x)$ and $r(x)$ exists and proving they are unique given each ordered pair of $\langle a(x), b(x)\rangle$.

- First we show such $q(x)$ and $r(x)$ exist. Define $\mathcal{S} \subseteq F[x]$ by

  $$\mathcal{S} = \{a(x) - b(x)r(x) : q(x) \in F[x]\}$$

  There will be polynomials with least degree. Pick one and we'll call it $r_0(x) = a(x) - b(x)q_0(x)$. Now we need to show $\deg(r_0) < \deg(b)$. Suppose not, i.e., $\deg(r_0) \geqslant \deg(b)$, and we expand $r_0(x)$ and $b(x)$ as

  $$\begin{aligned} r_0(x) &= \sum_{i=0}^{m} r_i x^i = r_m x^m + r_{m-1}x^{m-1} + \cdots + r_1 x + r_0 \\ b(x) &= \sum_{i=0}^{n} b_i x^i = b_n x^n + b_{n-1}x^{n-1} + \cdots + b_1 x + b_0 \end{aligned}$$

  with $r_m, b_n \neq 0$. Then $\deg(r_0) = m, \deg(b) = n$, and $m \geqslant n$. Since $b_n \neq 0$, by the definition of a field, both b_n and its reciprocal are in F, and since a field (or more generally, a ring) is closed under multiplication, $\dfrac{r_m}{b_n} \in F$ as well. Now consider the polynomial

  $$\begin{aligned} r_0(x) - \frac{r_m}{b_n}x^{m-n} \cdot b(x) &= a(x) - b(x)q_0(x) - \frac{r_m}{b_n}x^{m-n} \cdot b(x) \\ &= a(x) - b(x) \underbrace{\cdot \left[q_0(x) - \frac{r_m}{b_n}x^{m-n}\right]}_{\text{also a polynomial in } F[x]} \end{aligned}$$

  and we see that $r_0(x) - \dfrac{r_m}{b_n}x^{m-n} \cdot b(x)$ is in the set $\mathcal{S}$. Now rewrite the second term:

  $$\begin{aligned} \frac{r_m}{b_n}x^{m-n} \cdot b(x) &= \frac{r_m}{b_n}x^{m-n} \cdot (b_n x^n + b_{n-1}x^{n-1} + \cdots + b_1 x + b_0) \\ &= r_m x^m + \frac{r_m}{b_n}x^{m-n} \cdot (b_{n-1}x^{n-1} + \cdots + b_1 x + b_0) \\ &= r_m x^m + \sum_{i=0}^{n-1} \frac{r_m b_i}{b_n} x^{m-n+i} \end{aligned}$$

  Therefore the term $r_m x^m$ gets canceled in the polynomial $r_0(x) - \dfrac{r_m}{b_n}x^{m-n} \cdot b(x)$, and the degree of this new polynomial has degree strictly less than $m = \deg(r_0)$, contradicting the assumption that $r_0(x)$ has the lowest degree among all polynomials in $\mathcal{S}$. Therefore there exists a "remainder" polynomial $r(x)$ that satisfies $\deg(r) < \deg(b)$.

- Now for uniqueness. Suppose for $a(x), b(x) \in F[x]$ we have $q(x), r(x), q'(x), r'(x) \in F[x]$ such that

  $$\begin{aligned} a(x) &= b(x)q(x) + r(x) \qquad \deg(r) < \deg(b) \\ a(x) &= b(x)q'(x) + r'(x) \quad \deg(r') < \deg(b) \end{aligned}$$

Then

$$0_F = a(x) - a(x) = b(x) \cdot [q(x) - q'(x)] + [r(x) - r'(x)]$$
$$r(x) - r'(x) = b(x) \cdot [q'(x) - q(x)]$$

Since all of the polynomials involved are in a field (or more generally, an integral domain), by proposition 5.1.14 we have

$$\deg(r - r') = \deg(b) + \deg(q' - q)$$

By assumption we know $\deg(r) < \deg(b)$ and $\deg(r') < \deg(b)$. Therefore $\deg(r - r') < \deg(b)$. If the degree equation above holds true then $\deg(q' - q) < 0$. By proposition 5.1.12, the only possible case is when $\deg(q' - q) = -\infty$ and $q'(x) - q(x) = 0_F$. Then $\deg(r - r')$ must also be $-\infty$, meaning $r(x) = r'(x)$, which finishes the proof.

□

Problem 5.2.1 (5.2.2)**.** For each field F and each pair $a(x), b(x) \in F[x]$, find $q(x), r(x) \in F[x]$ such that $a(x) = b(x)q(x) + r(x)$ and $\deg(r) < \deg(b)$.

(1) $F = \mathbb{Q}, a(x) = x^5 + 2x^2 - 2, b(x) = x^3 + 7x + 1$.

(2) $F = \mathbb{Z}/13\mathbb{Z}, a(x) = x^3 + x^2 + 1b(x) = x + 11$.

Solution 5.2.1.

(1) $q(x) = x^2 + 7, r(x) = x^2 - 49x + 5$.

(2) $q(x) = x^2 - 10x + 6, r(x) = 0$. Note that since $a(x), b(x), q(x), r(x) \in \mathbb{Z}/13\mathbb{Z}$, we have $[-110x] = [6x]$ and $[65x] = [0]$. There are infinitely many ways to write $q(x)$ and $r(x)$.

$$\begin{array}{r|rrrrr}
 & & & +x^2 & & -7 \\ \cline{2-6}
x^3 +7x +1 & x^5 & & +2x^2 & & -2 \\
 & x^5 & +7x^3 & +x^2 & & \\ \cline{2-4}
 & & -7x^3 & +x^2 & & -2 \\
 & & -7x^3 & & -49x & -7 \\ \cline{3-6}
 & & & x^2 & -49x & +5
\end{array}
\qquad
\begin{array}{r|rrrr}
 & & +x^2 & -10x & +6 \\ \cline{2-5}
x +11 & x^3 & +x^2 & & +1 \\
 & x^3 & +11x^2 & & \\ \cline{2-3}
 & & -10x^2 & & +1 \\
 & & -10x^2 & -110x & \\ \cline{3-4}
 & & & +[6x] & +1 \\
 & & & +6x & +66 \\ \cline{4-5}
 & & & & +[0]
\end{array}$$

5.3 Euclid's Algorithm for Polynomials

Definition 5.3.1. A greatest common divisor $d(x) \in F[x]$ of $a(x), b(x) \in F[x]$ satisfies the following:

(1) $d(x)$ is a common divisor of $a(x)$ and $b(x)$, and

(2) if $f(x) \in F[x]$ is a common divisor of $a(x)$ and $b(x)$ then $f(x) \mid d(x)$.

Definition 5.3.2 (Euclid's Algorithm for Polynomials)**.** This is very similar to Euclid's Algorithm except that now we are dealing with two polynomials instead of two integers. Suppose F is a field and we have two nonzero polynomials $a(x), b(x) \in F[x]$. The following process is called Euclid's Algorithm for Polynomials:

$$\begin{aligned}
a(x) &= b(x)q_1(x) + r_1(x) && \deg(r_1) < \deg(b) \\
b(x) &= r_1(x)q_2(x) + r_2(x) && \deg(r_2) < \deg(r_1) \\
r_1(x) &= r_2(x)q_3(x) + r_3(x) && \deg(r_3) < \deg(r_2) \\
&\vdots \\
r_{n-2}(x) &= r_{n-1}(x)q_n(x) + r_n(x) && \deg(r_n) < \deg(r_{n-1}) \\
r_{n-1}x &= r_n(x)q_{n+1}(x) + r_{n+1}(x)
\end{aligned}$$

Since the degrees of the remainder polynomials is strictly decreasing, eventually it will reach $-\infty$, i.e., we will eventually come to a zero polynomial $r_{n+1}(x)$. The last nonzero polynomial remainder, $r_n(x)$, satisfies $r_n(x) = \gcd(a(x), b(x))$.

Remark. Greatest common divisors of polynomials are not unique. See the definition of associate below.

Definition 5.3.3. Nonzero polynomials $a(x), b(x) \in F[x]$ are 'me if $\gcd(a(x), b(x)) = 1_F$.

Definition 5.3.4. Two polynomials $a(x), b(x) \in F[x]$ are ***associate*** if there is a $\lambda \in F^\times$ such that $a(x) = \lambda \cdot b(x)$. The notation is $a(x) \sim b(x)$.

Example 5.3.1. Suppose $a(x) = x^5 + 2x^2 + 3x + 1$ and $b(x) = x^4 + 2x^3 + 4$ are members of $\mathbb{Z}/7\mathbb{Z}[x]$. By the algorithm we have

$$\begin{aligned}
x^5 + 2x^2 + 3x + 1 &= (x^4 + 2x^3 + 4)(x - 2) && +(4x^3 + 2x^2 - x + 2) \\
x^4 + 2x^3 + 4 &= (4x^3 + 2x^2 - x + 2)(2x + 3) && +(3x^2 - x + 5) \\
4x^3 + 2x^2 - x + 2 &= (3x^2 - x + 5)(-x + 5) && +(-5x + 5) \\
3x^2 - x + 5 &= (-5x + 5)(5x + 1)
\end{aligned}$$

The last nonzero remainder polynomial, $-5x+5$, is a GCD of $a(x)$ and $b(x)$. However it is not unique. Multiplying $-5x + 5$ with any elements of $(\mathbb{Z}/7\mathbb{Z})^\times$ yields other GCDs. For example,

$$4 \cdot (-5x + 5) = -20x + 20 = x - 1$$

is a GCD. Note that the multiplicative of 4 in $\mathbb{Z}/7\mathbb{Z}$ is 2. Suppose $a(x) = a'(x)(-5x + 5)$ and $b(x) = b'(x)(-5x + 5)$ then

$$a(x) = [2 \cdot a'(x)][4 \cdot (-5x + 5)] \text{ and } b(x) = [2 \cdot b'(x)][4 \cdot (-5x + 5)].$$

It is customary, though not essential, to make the GCD monic (with leading coefficient 1). Therefore in this example we say $\gcd(a(x), b(x)) = x - 1$.

Proposition 5.3.5. The relation $\sim$ is an equivalence relation on the set $F[x]$.

Proof.

- $\ast$ Reflexivity: $a(x) = 1_F \cdot a(x) \implies a(x) \sim a(x)$.
- $\ast$ Symmetry: If $a(x) \sim b(x)$ then there exists $\lambda \in F^\times$ such that $a(x) = \lambda \cdot b(x)$. Then λ^{-1}, the multiplicative inverse of λ, saatisfies $b(x) = \lambda^{-1} a(x)$ and thus $b(x) \sim a(x)$.
- $\ast$ Transitivity: if $a(x) \sim b(x)$ and $b(x) \sim c(x)$ then there exist $\lambda_1, \lambda_2 \in F^\times$ such that $a(x) = \lambda_1 b(x)$ and $b(x) = \lambda_2 c(x)$. Since $\lambda_1\lambda_2$ is in $F^\times$ as well, $a(x) = (\lambda_1\lambda_2) \cdot c(x)$ shows $a(x) \sim c(x)$.

□

Proposition 5.3.6. Given polynomials $a(x), b(x) \in F[x]$ we have

$$a(x) \sim b(x) \iff a(x) \mid b(x) \text{ and } b(x) \mid a(x)$$

Proposition 5.3.7. Suppose $a(x), b(x) \in F[x]$. If $d(x)$ is a GCD of $a(x)$ and $b(x)$, then so is every associate of $d(x)$.

Theorem 5.3.8. Given $a(x), b(x) \in F[x]$, $\gcd(a(x), b(x))$ can be expressed as a $F[x]$–linear combination of $a(x)$ and $b(x)$.

Proof. Similar to the proof of theorem 2.4.1.

Future reference: problem 5.3.1, proposition 5.4.5. □

Definition 5.3.9. A least common multiple $m(x) \in F[x]$ of $a(x), b(x) \in F[x]$ satisfies the following:

(1) $m(x)$ is a common multiple of $a(x)$ and $b(x)$, and

(2) if $f(x) \in F[x]$ is a common multiple of $a(x)$ and $b(x)$ then $m(x) \mid f(x)$.

Problem 5.3.1 (5.3.12)**.** Find polynomials $s(x), t(x) \in \mathbb{Z}/13\mathbb{Z}[x]$ satisfying

$$(6x^5 + x + 2) \cdot s(x) + (3x^4 - x^2 + 1) \cdot t(x) = 1.$$

or show no such polynomials exist.

Solution 5.3.1. First, we define $a(x) = 6x^5 + x + 2$ and $b(x) = 3x^4 - x^2 + 1$ and apply Euclid's Algorithm to $a(x)$ and $b(x)$:

$$\begin{aligned}
a(x) &= b(x)(2x) && +(2x^3 - x + 2) \\
b(x) &= (2x^3 - x + 2)(8x) && +(7x^2 - 3x + 1) \\
2x^3 - x + 2 &= (7x^2 - 3x + 1)(4x - 2) && +(2x + 4) \\
7x^2 - 3x + 1 &= (2x + 4)(10x + 11) && +9 \\
2x + 4 &= 9(6x - 1)
\end{aligned}$$

from which we conclude that 9 is a GCD of $a(x)$ and $b(x)$. Since $9 \cdot 3 = 27 = 1$ in $\mathbb{Z}/13\mathbb{Z}$, they are associate and 1 is a GCD of $a(x)$ and $b(x)$. Therefore by theorem 5.3.8 there exists $s(x)$ and $t(x)$ such that

$$a(x)s(x) + b(x)t(x) = 1.$$

Now rearranging each term in the algorithm above:

$$\begin{aligned}
2x^3 - x + 2 &= a(x) - (2x)b(x) \\
7x^2 - 3x + 1 &= b(x) - (8x)(2x^3 - x + 2) \\
&= b(x) - (8x)[a(x) - (2x)b(x)] \\
&= (-8x)a(x) + (16x^2 + 1)b(x) \\
&= (-8x)a(x) + (3x^2 + 1)b(x) \\
2x + 4 &= (2x^3 - x + 2) - (4x - 2)(7x^2 - 3x + 1) \\
&= a(x) - (2x)b(x) - (4x - 2)[(-8x)a(x) + (3x^2 + 1)b(x)] \\
&= (32x^2 - 16x + 1)a(x) + (-12x^3 + 6x^2 - 6x + 2)b(x) \\
&= (6x^2 - 3x + 1)a(x) + (x^3 + 6x^2 - 6x + 2)b(x) \\
9 &= 7x^2 - 3x + 1 - (10x + 11)(2x + 4) \\
&= (-8x)a(x) + (3x^2 + 1)b(x) - (10x + 11)[(6x^2 - 3x + 1)a(x) + (x^3 + 6x^2 - 6x + 2)b(x)] \\
&= (-60x^3 - 36x^2 + 15x - 11)a(x) + (-10x^4 - 71x^3 - 3x^2 + 46x - 21)b(x) \\
&= (5x^3 + 3x^2 + 2x + 2)a(x) + (3x^4 + 7x^3 - 3x^2 + 7x + 5)b(x) \\
1 &= 3 \cdot 9 \\
&= 3(5x^3 + 3x^2 + 2x + 2)a(x) + 3(3x^4 + 7x^3 - 3x^2 + 7x + 5)b(x) \\
&= (2x^3 - 4x^2 + 6x + 6)a(x) + (9x^4 - 5x^3 + 4x^2 - 5x + 2)b(x)
\end{aligned}$$

Therefore we've found a set of $s(x)$ and $t(x)$ that satisfy the equation $a(x)s(x) + b(x)t(x) = 1$:

$$\begin{cases} s(x) & = (2x^3 - 4x^2 + 6x + 6) \\ t(x) & = (9x^4 - 5x^3 + 4x^2 - 5x + 2) \end{cases}$$

5.4 Unique Factorization of Polynomials

Unless otherwise stated, let F be a field from now on.

Definition 5.4.1. Suppose $a(x) \in F[x]$ is a nonconstaant polynomial.

(a) $a(x)$ is said to be ***irreducible*** if for every factorization $a(x) = s(x)t(x)$ with $s(x), t(x) \in F[x]$, either $\deg(s)$ or $\deg(d) = 0$.

(b) $a(x)$ is said to be ***factorizable*** if there exist $s(x), t(x) \in F[x]$, both of which have degree > 0, such that $a(x) = s(x)t(x)$.

Remark. A nonconstant polynomial is either irreducible or factorizable. A constant polynomial, however, is said to be neither irreducible or factorizable.

Proposition 5.4.2. If $a(x) \in F[x]$ is irreducible and $b(x) \mid a(x)$, then either $b(x) \sim 1_F$ or $b(x) \sim a(x)$.

Proof. Since $b(x) \mid a(x)$, there exists a $c(x) \in F[x]$ such that $a(x) = b(x)c(x)$. Clearly neither $b(x)$ nor $c(x)$ are zero. However, since $a(x)$ is irreducible, either $b(x)$ or $c(x)$ has degree 0. If $b(x)$ has degree 0, then $b(x) = b(x)^{-1} \cdot 1_F$ which shows $b(x) \sim 1_F$. On the other hand, if $\deg(c) = 0$, then $c(x) \in F^\times$ and $a(x) \sim b(x)$. By symmetry, $b(x) \sim a(x)$.

Future reference: Fundamental Theorem of Arithmetic for Polynomials □

Proposition 5.4.3. Suppose $a(x), b(x) \in F[x]$, then $a(x)$ is irreducible $\iff$ $b(x)$ is irreducible.

Proof. We first show $\implies$. If $a(x)$ is irreducible then its divisors are either constant polynomials or its associaates. Since $a(x) \sim b(x)$ and $b(x) \sim a(x)$, there exists $\lambda \in F[x]$ such that $b(x) = \lambda \cdot a(x)$. Then we also have $a(x) = \lambda^{-1}b(x)$.

Suppose $b(x)$ is factorizable, i.e., there exist $s(x), t(x) \in F[x]$, both of which have degree > 0, such that $s(x)t(x) = b(x)$. Look at $a(x)$ again, which now equals $\lambda^{-1}s(x)t(x)$. Since $\deg(s)$ and $\deg(t)$ are both greater than 0, so are $\lambda^{-1}s(x)$ or $\lambda^{-1}t(x)$. Now we've found two ways to factorize $a(x)$ as the product of two non-constant polynomials — $a(x) = [\lambda^{-1}s(x)]t(x) = [\lambda^{-1}t(x)]s(x)$ — contradicting $a(x)$ being irreducible.

To show $\impliedby$, simply start by assuming $b(x)$ is irreducible. Then repeat the same process to show a contradiction. □

Proposition 5.4.4. Each nonzero polynomial $a(x) \in F[x]$ can be expressed as the product of irreducible polynomials.

Proof. We will approach this proof by strong induction. Let $\varphi(n)$ be the statement that each degree n polynomial can be factorized into irreducible polynomials.

- Clearly $\varphi(1)$ is true since a degree 1 polynomial can only be expressed as the product of one degree 1 and one degree 0 polynomial by proposition 5.1.14

- Now assume for some positive integer k, $\varphi(n)$ holds for all $n \in [1, k]$. Now look at degree $k+1$ polynomials. If this polynomial is irreducible then we are immediately done. If not, it can be factorized into polynomials with degrees greater than 0 but less than $k+1$. (Again, by proposition

5.1.14 since the two component polynomials both have degree greater than 0 and their sum is $k+1$.) By induction hypothesis, regarless of their degrees, both component polynomials can be factorized into a series of irreducible polynomials. Therefore all degree $k+1$ polynomials can be expressed as the product of some irreducible polynomials, hence $\varphi(k+1)$ is also true.

- Having proven both the base case and the inductive step, we are done with the main proof.

Future reference: Fundamental Theorem of Arithmetic for Polynomials □

Proposition 5.4.5. Suppose $a(x), b(x), p(x) \in F[x]$ with $p(x)$ irreducible. Then

$$p(x) \mid a(x)b(x) \implies p(x) \mid a(x) \text{ or } p(x) \mid b(x).$$

Proof. WLOG we look at $p(x)$ and $a(x)$ first. If $p(x)$ and its associates are the GCDs of $p(x)$ and $a(x)$, then we are immediately done with $p(x) \mid a(x)$.

Suppose $p(x)$ and its associates are not the GCDs, then $\gcd(p(x), a(x)) = 1_F$, i.e., $p(x)$ and $a(x)$ are co-prime. By theorem 5.3.8, there exist $s(x), t(x)$ such that $p(x)s(x) + a(x)t(x) = 1_F$. Therefore, multiplying both sides by $b(x)$ we have

$$p(x)b(x)s(x) + a(x)b(x)t(x) = b(x) \cdot 1_F = b(x).$$

Since $p(x) \mid p(x)$ (obviously) and $p(x) \mid a(x)b(x)$ as given, we conclude that $p(x) \mid b(x)$, which finishes the proof.

Future reference: proposition 5.4.6 □

Proposition 5.4.6. A stronger version of the previous proposition:

Suppose $p(x), a_1(x), a_2(x), \cdots, a_n(x) \in F[x]$ with $p(x)$ irreducible. Then

$$p(n) \mid \prod_{i=1}^{n} a_i(x) \implies p(n) \mid a_k(n) \text{ for some } 1 \leqslant k \leqslant n.$$

Proof. We approach this proof by (weak) induction. Let $\varphi(n)$ be the statement that if irreducible $p(n)$ divides the product of n irreducible polynomials then $p(n)$ divides at least one of these polynomials.

- $\varphi(1)$ is immediately true and $\varphi(2)$ is true as shown in proposition 5.4.5.

- Suppose $\varphi(k)$ is true and we want to show $\varphi(k+1)$ is true as well. Suppose we have $k+1$ irreducible polynomials, $a_1(x), a_2(x), \cdots, a_{k+1}(x)$. Their product, $\prod_{i=1}^{k+1} a_i(x)$, can be re-written as $a_1(x) \cdot \prod_{i=2}^{k+1} a_i(x)$. If $p(x) \mid a_1(x)$ then we are immediately done. If not, since $p(x) \mid a_1(x) \cdot \prod_{i=2}^{k+1} a_i(x)$, by proposition 5.4.6 it must be the case that $p(x) \mid \prod_{i=2}^{k+1} a_i(x)$. However, this term is a product of k irreducible polynomials, and by our induction hypothesis there exists at least one irreducible polynomial which $p(x)$ divides. Hence $\varphi(k) \implies \varphi(k+1)$.

- Having proven both the base cases and the indcutive step, we have therefore shown that $\varphi(n)$ holds true for all all $n \in \mathbb{Z}^+$. Hence $p(n) \mid \prod_{i=1}^{n} a_i(x) \implies p(n) \mid a_k(n)$ for some $1 \leqslant k \leqslant n$.

Future reference: Fundamental Theorem of Arithmetic for Polynomials □

Theorem 5.4.7 (Fundamental Theorem of Arithmetic for Polynomials)**.** Let $a(x) \in F[x]$ be aa nonconstant polynomial. Then there are irreducibles

$$p_1(x), p_2(x), \cdots, p_m(x) \in F[x]$$

whose product is $a(x)$. If $a(x) = \prod_{i=1}^{n} q_i(x)$ is another factorization of $a(x)$ into irreducibles then $m = n$ and, after reordering $q_1(x), q_2(x), \cdots, q_m(x)$ we have

$$p_i(x) \sim q_i(x) \text{ for all } 1 \leqslant i \leqslant m = n.$$

Proof. The existence of $p_1(x), p_2(x), \cdots, p_n(x)$ is proven in proposition 5.4.4. For the uniqueness of the factorization, suppose there are two ways to factorize $a(x)$ into irreducibles

$$p(x) = \prod_{i=1}^{m} p_i(x) = p_1(x)p_2(x)\cdots p_m(x)$$
$$q(x) = \prod_{i=1}^{n} q_i(x) = q_1(x)q_2(x)\cdots q_n(x)$$

such that $m \neq n$. WLOG assume $m \leqslant n$. Clearly $p_1(x) \mid p(x)$, so $p_1(x) \mid \prod_{i=1}^{n} q_i(x)$. By proposition 5.4.6 we know $p_1(x)$ dividers at least one irreducible $q-$polynomials. After reordering the $q_i(x)$'s we may assume that $p_1(x) \mid q_1(x)$. Since $q_1(x)$ is irreducible, either $p_1(x) \sim 1_F$ or $p_1(x) \sim q_1(x)$ by proposition 5.4.2. Clearly the first case is impossible since associates of 1_F are constant polynomails — not irreducibles. Therefore $p_1(x) \sim q_1(x)$, and there exists a $\lambda_1 \in F^\times$ such that $p_1(x) = \lambda_1 q_1(x)$. Therefore

$$a(x) = \prod_{i=1}^{m} p_i(x) = \lambda_1 q_1(x) \prod_{i=2}^{m} p_i(x) = q_1(x) \prod_{i=2}^{n} q_i(x)$$

Since F is an integral domain, so is $F[x]$ by theorem 5.1.16. Therefore cancellation law holds and we may cancel $q_1(x)$ from both sides:

$$\lambda_1 \prod_{i=2}^{m} p_i(x) = \prod_{i=2}^{n} q_i(x)$$

Look at $p_2(x)$ on the LHS. Similar to what we've previously done,$p_2(x)$ divides an irreducible $q-$polynomial from the RHS. After reordering, we may call this polynomail $q_2(x)$. Then $p_2(x) \sim q_2(x)$ and there exists a $\lambda_2 \in F^\times$ such that $p_2(x) = \lambda_2 q_2(x)$. Substitute $p_2(x)$ with $\lambda_2 q_2(x)$ and apply cancellation law again, we have

$$\lambda_1 \lambda_2 \prod_{i=3}^{m} p_i(x) = \prod_{i=3}^{n} q_i(x)$$

Repeat this process m times and we have

$$\prod_{i=1}^{m} \lambda_i = \prod_{i=m+1}^{n} q_i(x) \text{ and } p_i(x) \sim q_i(x) \text{ for all } 1 \leqslant i \leqslant m.$$

Look at the degree of both sides. Clearly the LHS is the product of degree 0 polynomials, so it has degree 0 aas well. If $n > m$ then the RHS is a nonempty product of polynomials of degree > 0, contradiction. Therefore $m = n$ and we have proven the Fundamental Theorem of Arithmetic for Polynomials.

Future reference: problem 5.5.1 □

Problem 5.4.1 (5.4.9)**.** Find two different ways to factor $x^2+x+8 \in \mathbb{Z}/10\mathbb{Z}[x]$ as a produict of monic degree one polynomials.

Solution 5.4.1. Observe that in $\mathbb{Z}/10\mathbb{Z}[x]$,

$$x^2 + x + 8 = \begin{cases} x^2 + x - 2 = (x-1)(x+2) \\ x^2 + x - 12 = (x-3)(x+4) \end{cases}$$

5.5 Roots of Polynomials

Definition 5.5.1. Suppose $a(x) \in F[x]$ and $r \in F$. We say r is a root of $a(x)$ if $a(r) = 0_F$.

Proposition 5.5.2. Suppose $a(x) \in F[x]$ and $r \in F$. Then

$$r \text{ is a root } \iff (x-r) \mid a(x)$$

Proof. We first show $\Longrightarrow$. Suppose $f \in F$ is a root of $a(x) \in F[x]$. We now define another polynomail $b(x) = x - r$. By the division algorithm for polynomials, there exist $f(x), g(x) \in F[x]$ such that

$$a(x) = f(x)b(x) + g(x) \text{ and } \deg(g) < \deg(b)$$

Now suppose $x = r$. The equation becomes

$$a(r) = f(r)b(r) + g(r) \implies 0 = 0 \cdot f(r) + g(r) \implies g(r) = 0$$

Therefore if r is a root then $a(x) = (x-r)f(x)$.

The other direction, $\Longleftarrow$, is obvious. If $(x-r) \mid a(x)$ then there exists $f(x)$ such that $a(x) = (x-r)f(x)$. If $x = r$ then $a(x) = 0 \cdot f(x) = 0$. Thus r is a root. □

Proposition 5.5.3. A polynomial $f(x) \in F[x]$ of degree n has at most n distinct roots in F.

Proof. We approach this by induction. Let $\varphi(n)$ be the statement that a polynomial f of degree n on a field F has at most n distinct roots in F.

- ⋆ Clearly every degree 1 polynomial has one root. Therefore $\varphi(1)$ is true.
- ⋆ Suppose $\varphi(k)$ is true, i.e., every degree k polynomial has at most k distinct roots. Now look at degree $k+1$ polynomials. If $a(x)$ is of degree $k+1$ and r is a root of $a(x)$, then $a(x) = (a-r)b(x)$ for some degree k polynomial $b(x)$. By our induction hypothesis $b(x)$ has at most k distinct roots. Since $a(x) = 0$ if and only if $a = r$ or $b(x) = 0$, $a(x)$ has at most $k+1$ roots — one from $a = r$ and k from roots of $b(x)$. Hnece inductive step proven.
- ⋆ Having proved both the base case and the inductive step, we conclude that $\varphi(n)$ holds for all $n \in \mathbb{Z}^+$, i.e., a polynomial $f(x) \in F[x]$ of degree n has at most n distinct roots in F.

□

Proposition 5.5.4. Suppose $a(x) \in F[x]$.

(a) If $\deg(a) = 1$ then $a(x)$ is irreducible.

(b) If $\deg(a) = 2$ or $\deg(a) = 3$ then $a(x)$ is irreducible $\iff$ it has no roots.

Proof. Part (a) is obvious, and we will only look at part (b).

We first show $\Longrightarrow$. If $a(x)$ is irreducible then $(x-r)$ does not divide $a(x)$ for any $r \in F$, which implies $a(x)$ has no roots.

Now look at $\Longleftarrow$. if $a(x)$ has no roots, then it does not have any factor of degree 1. Suppose $a(x) = b(x)c(x)$. Since F is an integral domain, $\deg(a) = \deg(b) + \deg(c)$ by proposition 5.1.14. If

$\deg(a) = 2$ then the degrees of $b(x)$ and $c(x)$ can only be 0 and 2 or 2 and 0. If $\deg(a) = 3$ then the degrees of $b(x)$ and $c(x)$ can only be 0 and 3 or 3 and 0. In all these cases, either $b(x)$ or $c(x)$ has degree 0, and thus $a(x)$ is irreducible.

Future reference: problem 5.5.1 □

Remark. Look at the $\Longleftarrow$ part again. If $\deg(a) > 3$, then it is possible to factor a degree 4 polynomial into two degree 2 irreducibles. In this case $a(x)$ would be factorizable but still without roots. For example, $x^4 + 3x^2 + 2 \in (\mathbb{Z}/7\mathbb{Z})[x]$ can be factorized as $(x^2 + 1)(x^2 + 2)$, both of which hare irreducible.

Theorem 5.5.5 (Rational Root Test)**.** Suppose a polynomial

$$a(x) = \sum_{i=0}^{n} a_i x_i = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 \in \mathbb{Z}[x]$$

with $a_n \neq 0$ has a rational root $\dfrac{p}{q}$ where $p \in \mathbb{Z}, q \in \mathbb{Z}^+$, and $\gcd(p, q) = 1$. Then $p \mid a_0$ and $q \mid a_n$.

Proof. Plugging in $x = \dfrac{p}{q}$ to the polynomial yields

$$0 = a_n \left(\frac{p}{q}\right)^n + a_{n-1} \left(\frac{p}{q}\right)^{n-1} + \cdots + a_1 \left(\frac{p}{q}\right) + a_0$$

and multiplying both sides by q^n yields

$$0 = a_n p^n + a_{n-1} p^{n-1} q + \cdots + a_1 p q^{n-1} + a_0 q^n$$

Clearly q divides the LHS and the sum of all but the first term on the RHS. Therefore $q \mid a_n p^n$. Since $\gcd(p, q) = 1$, it follows that $q \mid a_n$.

Similarly, p divides the LHS and the sum of all but the last term on the RHS. Therefore $p \mid a_0 q^n$. Since $\gcd(p, q) = 1$, it follows that $p \mid a_0$. □

Theorem 5.5.6 (Quadratic Formula)**.** Suppose $f(x) = ax^2 + bx + c \in F[x]$ with $2a \in F^\times$, and set $\Delta = b^2 - 4ac$.

(a) If there is a $\delta \in F$ such that $\delta^2 = \Delta$, then the roots of $f(x)$ are $\dfrac{-b \pm \delta}{2a}$.

(b) If there is no such $\delta \in F$, then $f(x)$ has no roots in F.

Proof. The proof is done by completing the square. We focus on $f(x)$ for now.

$$\begin{aligned} f(x) &= ax^2 + bx + c \\ &= ax^2 + 2 \cdot a \cdot \frac{b}{2a} x + c \\ &= ax^2 + 2 \cdot a \cdot \frac{b}{2a} x + \frac{b^2}{4a} + \left(c - \frac{b^2}{4a}\right) \\ &= a\left(x^2 + 2 \cdot \frac{b}{2a} x + \frac{b^2}{4a^2}\right) + \left(c - \frac{b^2}{4a}\right) \\ &= a\left(x + \frac{b}{2a}\right)^2 + \left(c - \frac{b^2}{4a}\right) \end{aligned}$$

If $f(x) = 0$, then $-a\left(x + \dfrac{b}{2a}\right)^2 = c - \dfrac{b^2}{4a}$. Arranging the terms gives

$$\left(x + \frac{b}{2a}\right)^2 = -\frac{1}{a}\left(c - \frac{b^2}{4a}\right) = \frac{b^2 - 4ac}{4a^2}$$

We've set $\Delta = b^2 - 4ac$. If there exists $\delta \in F$ such that $\delta^2 = \Delta$, then $\delta = \sqrt{b^2 - 4ac}$. Therefore

$$\left(x + \frac{b}{2a}\right) = \sqrt{\frac{b^2 - 4ac}{4a^2}} = \pm\frac{\delta}{2a} \implies x = -\frac{b}{2a} \pm \frac{\delta}{2a} = \frac{-b \pm \delta}{2a}.$$

Note that we need the prerequisite that $2 \neq 0$ in F, i.e., char $F \neq 2$ or $2a \in F^\times$. Only by doing so can we guarentee that the division with denominator $2a$ is defined.

If there is no such δ then we cannot find a root $f(x)$ in F.

Future reference: problem 5.5.1 □

Problem 5.5.1 (5.5.9-14)**.** Factor x^4+1 into irreducible factors in $\mathbb{Q}[x], \mathbb{R}[x], \mathbb{C}[x]$, $\mathbb{Z}/2\mathbb{Z}[x], \mathbb{Z}/3\mathbb{Z}[x]$, and $\mathbb{Z}/5\mathbb{Z}[x]$.

Solution 5.5.1. By completing the square, we have

$$\begin{aligned} x^4 + 1 &= (x^4 + 2x^2 + 1) - 2x^2 \\ &= (x^2 + 1)^2 - 2x^2 \\ &= (x^2 + \sqrt{2}x + 1)(x^2 - \sqrt{2}x + 1) \end{aligned}$$

which is a factorization of x^4+1 in $\mathbb{R}[x]$ into two degree 2 polynomials. By the Quadratic Formula, both have negative determinants, which, by proposition 5.5.4, implies that both are irreducible. Then the Fundamental Theorem of Arithmetic for Polynomials implies that $x^4 + 1$ cannot be further factorized in $\mathbb{Q}[x]$ (since the irreducible polynomials in $\mathbb{R}[x]$ include both irrational and rational coefficients).

Now look at factorization over $\mathbb{C}$. Re-write $x^4 + 1 = 0$ as $x^4 = -1$. We know that $-1 = e^{i\pi}$. By DeMoivre's Theorem, $x^4 = -1$ has 4 distinct roots:

$$x_1 = e^{\pi i/4}, x_2 = e^{3\pi i/4}, x_3 = e^{5\pi i/4}, x_4 = e^{7\pi i/4}$$

which implies

$$x^4 + 1 = (x - e^{\pi i/4})(x - e^{3\pi i/4})(x - e^{5\pi i/4})(x - e^{7\pi i/4}) \in \mathbb{C}[x]$$

For $(\mathbb{Z}/2\mathbb{Z})[x]$, note that $x^4 + 1 = x^4 - 1$ and $x^2 + 1 = x^2 - 1$. Therefore

$$\begin{aligned} x^4 + 1 = x^4 - 1 &= (x^2 + 1)(x^2 - 1) \\ &= (x^2 - 1)^2 = (x + 1)^2(x - 1)^2 \\ &= (x + 1)^4 \in (\mathbb{Z}/2\mathbb{Z})[x] \end{aligned}$$

For $\mathbb{Z}/3\mathbb{Z}[x]$, note that $x^4 + 1 = x^4 - 3x^2 + 1$. Therefore

$$\begin{aligned} x^4 + 1 &= x^4 - 3x^2 + 1 \\ &= (x^4 - 2x^2 + 1) - x^2 \\ &= (x^2 - 1)^2 - x^2 \\ &= (x^2 - x - 1)(x^2 + x - 1) \in \mathbb{Z}/3\mathbb{Z}[x] \end{aligned}$$

A quick examination shows no $\delta \in \mathbb{Z}/3\mathbb{Z}$ satisfies $\delta^2 = 2$, the value of Δ for both polynomials. Therefore we've factored $x^4 + 1$ into irreducibles in $\mathbb{Z}/3\mathbb{Z}[x]$.

Now for $\mathbb{Z}/5\mathbb{Z}[x]$. Note that $x^4 + 1 = x^4$. Therefore

$$x^4 + 1 = x^4 - 4 = (x^2 + 2)(x^2 + 3) \in \mathbb{Z}/5\mathbb{Z}[x]$$

A quick examination shows both polynomials have no root, and by proposition 5.5.4 they are irreducible, so we've found the desired factorization of $x^4 + 1$.

5.6 Derivatives and Multiple Roots

Definition 5.6.1. Suppose a polynomial

$$f(x) = \sum_{i=0}^{n} = a_i x^i = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0.$$

Then its ***derivative***, $f'(x)$, can be expressed as

$$f'(x) = \sum_{i=1}^{n} i a_i x^{i-1} = n a_n x^{n-1} + (n-1) a_{n-1} x^{n-2} + \cdots + 2a_2 x + a_1.$$

Remark. It is always true that $\deg(f') < \deg(f)$, since the degree n term exists in $f(x)$ but not $f'(x)$. However, it is not always true that $\deg(f') = \deg(f) - 1$. For example, a degree n polynomial in $\mathbb{Z}/n\mathbb{Z}[x]$ has derivative of degree no more than $n-2$ since the degree $n-1$ term, $na_n x^{n-1}$, equals 0 since $n = 0$ in $\mathbb{Z}/n\mathbb{Z}$. Nevertheless, if $f(x) \in \mathbb{Q}[x], \mathbb{R}[x]$, or $\mathbb{C}[x]$, the equation $\deg(f') = \deg(f) - 1$ is always true.

Proposition 5.6.2. Suppose $f(x)$ and $g(x)$ are polynomials. Then

(a) $(f+g)' = f' + g'$

(b) $(f \cdot g)' = (f')g + f(g')$.

Definition 5.6.3. Suppose $f(x) \in F[x]$ and $\alpha \in F$ is a root of $f(x)$.

(a) α is a ***root of multiplicity*** n if $(x-\alpha)^n$ divides $f(x)$ but $(x-\alpha)^{n+1}$ doesn't.

(b) A root of multiplicity 1 and 2 are called ***simple root*** and ***double root***, respectively.

(c) A root is either a simple root or a ***multiple root***.

Remark. Given $f(x) \in F[x]$, the following are equivalent:

(a) α is a root of multiplicity n.

(b) There exists a $g(x) \in F[x]$ such that $f(x) = (x-\alpha)^n g(x)$ and $g(\alpha) \neq 0$.

Proposition 5.6.4. Suppose $f(x) \in F[x]$ has multiple root $\alpha \in F$. Then $\alpha \mid \gcd(f, f')$.

Proof. If α is a multiple root of $f(x)$, then there exists $g(x) \in F[x]$ such that $f(x) = (x-\alpha)^2 g(x)$. By product rule,

$$\begin{aligned} f'(x) &= 2(x-\alpha)g(x) + (x-\alpha)^2 g'(x) \\ &= (x-\alpha)\left(2g(x) + (x-\alpha)g'(x)\right) \end{aligned}$$

which implies $(x-\alpha)$ divides $f'(x)$ as well. □

5.7 Gauss's Lemma and Eisenstein's Criterion

Definition 5.7.1. A polynomial $f(x) = \sum_{i=0}^{n} = a_i x^i = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 \in \mathbb{Z}[x]$ is ***primitive*** if the GCD of *all* coefficients is 1. (The coefficients do not necessarily need to be pairwise co-prime.)

Proposition 5.7.2. Every nonzero polynomial $f(x) \in \mathbb{Q}[x]$ can be written as the product of a nonzero rational number times a primitive polynomial in $\mathbb{Z}[x]$.

Proof. I will simply show by an example here. The proof follows exactly the same way. Consider

$$f(x) = \frac{4}{5}x^2 + \frac{2}{3}x + \frac{2}{5}.$$

Multiplying both sides by 3 × 5 we have

$$15f(x) = 12x^2 + 10x + 6 = 2(6x^2 + 5x + 3).$$

Therefore $f(x) = \frac{2}{15}(6x^2 + 5x + 3)$.

Future reference: Gauss's Lemma II □

Theorem 5.7.3 (Gauss's Lemma I)**.** Suppose $f(x), g(x) \in \mathbb{Z}[x]$ are primitive polynomials Then the product $f(x)g(x)$ is also primitive.

Proof. Suppose not, then there exists some integer $d > 1$ such that d divides all coefficients of the polynomial $f(x)g(x)$. Since $d > 1$, there exists a prime number p that divides d. It follows that p divides all coefficients of $f(x)g(x)$ as well. Now for $a(x) \in \mathbb{Z}[x]$, define $\bar{a}(x) \in \mathbb{Z}/p\mathbb{Z}[x]$ as the polynomial obtained by reducing all coefficients of $a(x)$ modulo p. For example,

$$a(x) = 7x^3 + 10x^2 - 11x + 2 \in \mathbb{Z}[x] \xrightarrow{p=5} \bar{a}(x) = 2x^3 + 4x + 2 \in \mathbb{Z}/5\mathbb{Z}[x].$$

Now, since all coefficients of $f(x)g(x)$ are multiples of p, we have $\bar{f}(x)\bar{g}(x) = 0$. By proposition 5.1.9, $\mathbb{Z}/p\mathbb{Z}$ is an integral domain, and so is $\mathbb{Z}/p\mathbb{Z}[x]$ by theorem 5.1.16. Therefore either $\bar{f}(x) = 0$ or $\bar{g}(x) = 0$. In either case we have one polynomial whose coefficients are all divisible by p, contradicting $f(x)$ and $g(x)$ both being primitive.

Future reference: Gauss's Lemma II, Eisenstein's Criterion □

Theorem 5.7.4 (Gauss's Lemma II)**.** Suppose $f(x) \in \mathbb{Z}[x]$ is a nonconstant polynomial that can be factored as $f(x) = a(x)b(x)$ where $a(x), b(x) \in \mathbb{Q}[x]$. Then there exist $c(x), d(x) \in \mathbb{Z}[x]$ such that

$$f(x) = c(x)d(x) \text{ and } c(x) \sim a(x), d(x) \sim b(x)$$

Proof. By proposition 5.7.2, there exist primitive $A(x), B(x) \in \mathbb{Z}[x]$ and $s, t \in \mathbb{Q}$ such that $a(x) = sA(x)$ and $b(x) = tB(x)$. Therefore,

$$f(x) = a(x)b(x) = stA(x)B(x).$$

Clearly $st \in \mathbb{Q}$, and we will show in a moment that $st \in \mathbb{Z}$ as well. Now suppose $st = \frac{p}{q}$ where $p \in \mathbb{Z}$ and $q \in \mathbb{Z}^+$. Then,

$$f(x) = \frac{p}{q}A(x)B(x) \implies qf(x) = pA(x)B(x)$$

Look at the RHS. By Gauss's Lemma I we know that since $A(x)$ and $B(x)$ are both primitive, so is $A(x)B(x)$. Therefore the GCD of all coefficients of $pA(x)B(x)$ is just p. Now look at the LHS. Suppose the GCD of all coefficients of $f(x)$ is d. Then $qd = p$ which implies $q \mid p$, and $st = \frac{p}{q}$ is indeed an integer. Now simply take

$$c(x) = stA(x) = ta(x) \text{ and } d(x) = B(x) = \frac{1}{t}b(x)$$

and we are done.

Future reference: Eisenstein's Criterion □

Theorem 5.7.5 (Eisenstein's Criterion)**.** Suppose polynomial

$$a(x) = \sum_{i=0}^{n} a_i x^i = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 \in \mathbb{Q}[x]$$

has integer coefficients, and there is a prime p satisfying

(a) p divides all coefficients but a_n, and

(b) p^2 does not a_0,

then $a(x)$ is irreducible (in $\mathbb{Q}[x]$).

Proof. We will prove by contradiction. Suppose for a factorizable $a(x)$ there exists such p satisfying both (a) and (b) listed in the theorem. Then there exist $b(x), c(x) \in \mathbb{Q}[x]$ such that

$$a(x) = b(x)c(x) \text{ and } 0 < \deg(b_o) < \deg(a), 0 < \deg(c) < \deg(a).$$

By Gauss's Lemma II, we may simply assume $b(x), c(x) \in \mathbb{Z}[x]$. Then we can expand them as

$$b(x) = \sum_{i=0}^{s} b_i x^i = b_s x^s + b_{s-1}x^{s-1} + \cdots + b_1 x + b_0$$
$$c(x) = \sum_{i=0}^{t} c_i x^i = c_t x^t + c_{t-1}x^{t-1} + \cdots + c_1 x + c_0$$

Then, $s, t > 0$, $s + t = n$ (granted by proposition 5.1.14),$a_0 = b_0 c_0$, and $a_n = b_s c_t$.

Since $p \mid a_0$ but $p^2 \nmid a_0$, it follows that, between b_0 and c_0, one is divisible by p while the other is not. WLOG assume $p \nmid b_0$. Also, since $p \nmid a_n$, it follows that $p \nmid b_s$. Recall the way how we define $\bar{a}(x)$ given $a(x)$ in Gauss's Lemma I. We will use it here again. Now look at $\bar{b}(x) \in \mathbb{Z}/p\mathbb{Z}[x]$.

Clearly, $\bar{b}(x)$ is nonconstant since $p \nmid b_s$ and the term $b_s x^s$ is nonzero. Also, x does not divide $\bar{b}(x)$ since $p \nmid b_0$ and the term b_0 is nonzero as well. Let $\bar{q}(x) \in \mathbb{Z}/p\mathbb{Z}[x]$ be an irreducible divisor of $\bar{b}(x)$. Since $x \nmid \bar{b}(x)$ as previously shown, we know $\bar{q}(x) \nsim x$ (easily proven by contradiction).

Now back to $a(x)$ and $\bar{a}(x)$. Since $\bar{a}(x) = \bar{b}(x)\bar{c}(x)$, $\bar{q}(x)$ divides $\bar{b}(x)$ implies $\bar{q}(x) \mid \bar{a}(x)$. However, since all coefficients of $a(x)$ but a_n are divisible of p, we have $\bar{a}(x) = a_n x^n$. Does $a_n x^n$ have any irreducible factors that are not associates of x? A quick examination suggests no, which leads to a contradiction. Hence we are done proving Eisenstein's Criterion. □

Remark. We say an irreducible polynomial $a(x) \in \mathbb{Q}[x]$ with integer coefficients is "Eisenstein at p" if p is a prime that meets the Eisenstein Criterion.

Problem 5.7.1 (5.7.7)**.** Suppose p is a prime.

(a) Show that $\dfrac{[(x+1)^p - 1]}{x} \in \mathbb{Q}[x]$ is irreducible.

(b) Use part (a) to prove that

$$\frac{x^p - 1}{x - 1} = \sum_{i=0}^{p-1} x^i = x^{p-1} + x^{p-2} + \cdots + x + 1 \in \mathbb{Q}[x]$$

is irreducible.

Solution 5.7.1.

(a) Applying Binomial Theorem yields

$$\begin{aligned}\frac{[(x+1)^p - 1]}{x} &= \frac{1}{x} \cdot \left(\sum_{i=0}^{p} \binom{p}{i} x^i - 1 \right) \\ &= \frac{1}{x} \cdot \left[\binom{p}{0} x^p + \binom{p}{1} x^{p-1} + \cdots + \binom{p}{p-1} x + \binom{p}{p} 1 - 1 \right] \\ &= \frac{1}{x} \cdot \left[\binom{p}{0} x^p + \binom{p}{1} x^{p-1} + \cdots + \binom{p}{p-1} x \right] \\ &= \binom{p}{0} x^{p-1} + \binom{p}{1} x^{p-2} + \cdots + \binom{p}{p-2} x + \binom{p}{p-1} \\ &= x^{p-1} + \binom{p}{1} x^{p-2} + \cdots + \binom{p}{p-2} x + p\end{aligned}$$

which is Eisenstein at p. Hence proven. □

(b) Define $f(x) = \dfrac{x^p - 1}{x - 1} = \sum_{i=0}^{p-1} x^i$. If $f(x)$ is factorizable, then there exist $g(x), h(x) \in \mathbb{Q}[x]$ such that $f(x) = g(x)h(x)$. It follows that $f(x+1) = g(x+1)h(x+1)$. Therefore we know that if $f(x)$ is factorizable then so is $f(x+1)$. However, $f(x+1)$ is exactly the fraction in part (a) which has been proven irreducible. Therefore $f(x)$ is irreducible as well. □

Problem 5.7.2 (5.7.8)**.** Show that there are infinitely many integers k such that $x^4 + 2x^2 + k$ is irreducible in $\mathbb{Q}[x]$.

Solution 5.7.2. If we can make $x^4 + 2x^2 + k$ Eisenstein at 2 then we are immediately done. To do so, k has to be a multiple of 2 but not 4. Since the set

$$\{2a + 2 : a \in \mathbb{Z}\}$$

has infinitely many (distinct) elements, each of which is a multiple of 2 but not 4, we are done with the proof. □