Olympiad Number Theory: An Abstract Perspective

Thomas J. Mildorf

June 6, 2010

Abstract. This paper develops some basic theorems in elementary number theory, especially those appearing on mathematical Olympiads, from the perspective of groups, rings, and group actions.

1 Introduction

In the mid seventeenth century, the French lawyer Pierre de Fermat wrote a letter in which he challenged his friend Frénicle de Bessy to prove a famous little theorem [1]. A secretive purveyor of mathematical puzzles in his day, Fermat is credited with the discovery of many early results underlying modern calculus, probability theory, and diophantine analysis [2]. He is remembered in particular for a pair of theorems, his Little and Last theorems. The latter withstood 357 years of investigation before a sophisticated solution was found (finally!) by Andrew Wiles in 1994 [3]. The former, which is elementary, will be proved in this paper. Indeed, Fermat's Little Theorem is a central piece of our discussion, which develops elementary number theory and modular arithmetic from the perspective of abstract algebra.

We focus on the groups and rings of algebra. Groups, having the existence of unique inverses as their chief property, capture the essence of multiplication in modular arithmetic. Rings simply incorporate the addition operation. With this perspective, emphasizing abstraction, we suggest to the reader that results such as Fermat's Little Theorem arise naturally as a progression that characterizes multiplicative order with increasing specificity.

We also expend considerable energy motivating the discussion. Bystanders see mathematics as a dull collection of axioms. Mathematicians, on the other hand, have been known compare it to a Big Game Hunt in the Sahara Desert. Thus, our second idea is to impart a good sense of the behavior behind these axioms: what they do and how they relate.

To this end, the organization is as follows. In Section 2, we give basic definitions, including all of those that will be assumed later. It may be skipped if the reader feels comfortable with groups; for students making this judgement, the notions of subgroups, cosets, quotient groups, and group homomorphisms

should not be especially encumbering. In Section 3, we introduce basic elements of group actions and ring theory. We start with the general Orbit-Stabilizer Theorem. From it, we arrive easily at Lagrange's Theorem, Euler's Theorem, Fermat's Little Theorem, Wilson's Theorem, and Eisenstein's Criterion. Section 4 focuses on finding elements with specific orders. Using Fermat's Little Theorem, we connect the discussion with multiplication in the field $\mathbb{Z}/p\mathbb{Z}$. We then build up the precise structure of $\mathbb{Z}/n\mathbb{Z}$, proving and using the Chinese Remainder Theorem in the process. At last, we establish the Primitive Root Theorem. As a bonus, we include an additional section, which introduces the *combinatorial Nullstellensatz,* a powerful theorem concerning the zeros of multivariate polynomials. With it, one can solve many problems by writing down and analyzing a suitably constructed polynomial.

2 Fundamentals

As an aid to the reader, we provide in this section a sketch of the results arising in the first week or so of an introductory course in abstract algebra. The propositions and their corollaries in this section are manifold, though simple; we leave some of their proofs as exercises for the reader.

Definition 2.1. A *group* $(G, \cdot)$ is a set G together with a binary law of composition $\cdot : G \times G \to G$ having the following three properties.

1. Associativity: the equality $a \cdot (b \cdot c) = (a \cdot b) \cdot c$ holds for all $a, b, c \in G$.
2. Identity: an element $e \in G$ exists, such that $g \cdot e = g = e \cdot g$ for all $g \in G$.
3. Inverses: for each element $g \in G$ there exists an $h \in G$ such that $g \cdot h = e$.

A group is called *abelian* or *commutative* if the law of composition is commutative.

For example, the integers under addition, or the positive reals under multiplication are examples of groups. The underlying set need not be infinite. For example, let n be a positive integer and define $X = \{1, 2, \ldots n\}$. There are $n!$ permutations of X, or bijective functions $\sigma : X \to X$. Under the natural composition of maps, these permutations give rise to the *symmetric group* S_n.

The immediate consequences of the group axioms are numerous; we collect several here.

Proposition 2.1. *A group has a unique identity element.*

Proof. Suppose that e and f are identity elements. Then $e = e \cdot f = f$. □

Proposition 2.2 (Cancellation Law)**.** *If $a, b, g \in G$ are such that $a \cdot g = b \cdot g$, then $a = b$.*

Proof. Multiply through by g^{-1}, obtaining

$$a = a \cdot e = a \cdot (g \cdot g^{-1}) = (a \cdot g) \cdot g^{-1} = (b \cdot g) \cdot g^{-1} = b \cdot (g \cdot g^{-1}) = b \cdot e = b.$$

□

Proposition 2.3 (Mutuality)**.** *If $g \cdot h = e$, then $h \cdot g = e$ as well.*

Proof. Write

$$e \cdot h = h \cdot e = h \cdot (g \cdot h) = (h \cdot g) \cdot h.$$

Now apply the Cancellation Law. □

Proposition 2.4. *Each group element $g \in G$ has a unique inverse.*

Proof. Let $g \in G$ and suppose that $x, y \in G$ are such that $g \cdot x = g \cdot y = e$. By Mutuality, $x \cdot g = e$, so that

$$x = x \cdot e = x \cdot (g \cdot y) = (x \cdot g) \cdot y = e \cdot y = y.$$

□

To promote diligence, we shall leave the proofs of the remaining propositions in this section as exercises.

It is customary to exchange $\cdot$ for $+$ as the notation for composition in an abelian group. Respecting this convention, the identity of an arbitrary group is often denoted 1, while the identity element is denoted 0 in a commutative group. Considering its uniqueness, the inverse of an element $g \in G$ is denoted g^{-1}; under this notation, mutuality becomes $(g^{-1})^{-1} = g$. As is typical, we will henceforth omit the operator when context makes it clear what composition is meant.

The associativity condition makes finite products of group elements unambiguous, and the two-sidedness of identity elements makes them and inverses unique. Essentially, the group axioms allow us to make sense of so-called words of a group, such as $abca^{-1}db^{-1}a$, as group elements. In particular, expressions such as $g^3 = ggg$ and $g^{-4} = g^{-1}g^{-1}g^{-1}g^{-1}$ have natural interpretations.

Definition 2.2. A *group homomorphism* between the groups G and H is a map $\phi : G \to H$ such that $\phi(ab) = \phi(a)\phi(b)$ for all $a, b \in G$. A homomorphism has as *kernel*, denoted $\ker(\phi) := \{g \mid g \in G, \phi(g) = e_H\}$, and an *image*, denoted $\mathrm{im}(\phi) := \{\phi(g) \mid g \in G\}$.

Proposition 2.5 (Composition)**.** *Suppose that $\phi : G \to H$ and $\varphi : H \to K$ are group homomorphisms. Then the composition $\varphi \circ \phi : G \to K$ is also a group homomorphism.*

Homomorphisms often have other properties. For instance, a homomorphism $\phi : G \to H$ is *injective* if $\ker(\phi) = \{e_G\}$ and *surjective* if $\mathrm{im}(\phi) = H$. are *isomorphic,* denoted $G \cong H$. A homomorphism that is both injective and surjective is called an *isomorphism.* A trivial example is the identity map $\nVdash_G : G \to G$ is an isomorphism of G.

Proposition 2.6. *Let $\phi : G \to H$ be a group homomorphism. Then ϕ is an isomorphism if and only if its inverse is.*

Thus, two groups G and H are called *isomorphic* if there exists an isomorphism between them. This property, denoted $G \cong H$, defines equivalence classes of groups because the identity map gives reflexivity, the preceding gives symmetry, and composition gives transitivity. Isomorphisms are considered strong equivalence relations, and one is usually only interested in establishing results "up to isomorphism." Indeed, the task of classifying finite simple groups means grouping them into equivalence classes by isomorphism.

Definition 2.3. The *order* of a group $(G, \cdot)$ is the cardinality of the set G and is denoted by $|G|$. If $a \in G$ and there exists a positive integer n such that $a^n = e$, then we say that a has *order* n where n is the smallest such integer.

For instance, one speaks of *the* finite group of order two comprises because all groups of two elements are isomorphic. If $G = \{1, g\}$, then $g \cdot g = 1$ because one also requires $g \cdot 1 = g$ and $g \cdot g \neq g \cdot 1$. Then an isomorphism $\phi : G \to G'$ between two such groups is given by $\phi(1) = 1'$ and $\phi(g) = g'$.

Definition 2.4. A subset H of G is called a *subgroup* of group G if composition restricts to a map $H \times H \to H$ that makes H into a group. A subgroup N of G is called normal if it is the kernel of a homomorphism from G, that is, if there exists some group homomorphism $\phi : G \to K$ such that $N = \ker(\phi)$.

A subgroup inherits the identity element and inverses from the overlying group.

The notation for the product of two group elements is often extended to the following products of an element and a subset:

$$H \subset G \text{ and } a \in G, \text{ then } aH := \{ah \,|\, h \in H\} \text{ and } Ha := \{ha \,|\, h \in H\}.$$

Definition 2.5. If H is a subgroup of G, then a *left coset* of H is a set of the form aH for some $a \in G$, and a *right coset* is a set Ha. The *index* of a subgroup H is the number of distinct left cosets of H in G, denoted $[G : H]$.

Definition 2.6. A subgroup N of G is *normal* if it is invariant under conjugation, or $gNg^{-1} = N$ for all $g \in G$. If N is a normal subgroup of G, then the *quotient group* G/N is the set of left cosets of N with the law of composition $(gN)(hN) = (gh)N$ and having order $[G : N]$.

Lemma 2.7 (First Isomorphism Theorem)**.** *If $\phi : G \to H$ is a group homomorphism, then*

$$G/\ker(\phi) \cong im(\phi).$$

Lemma 2.8. *Suppose that H and K are normal subgroups of G with $H \subseteq K$. Then K/H is a normal subgroup of G/H and*

$$(G/H)/(K/H) \cong G/K.$$

Exercises

Begin the following problems assuming that G is a group.

1. Verify that the rational numbers are a group under addition but not under multiplication. What are some subsets of the rational numbers that constitute a group under mutliplication?

2. Prove that for any $a \in G$, the set $\langle a \rangle = \{a^n : n \in \mathbb{Z}\}$ is a subgroup of G. This subgroup is called the *cyclic subgroup of G generated by a*. If G is a group and an element $g \in G$ is such that $\langle g \rangle = G$, then the group G is *cyclic* and g is a *generator*.

3. A subset H of a group G is a subgroup of G iff for all $a, b \in H$, the element ab^{-1} is in H.

4. The intersection of two subgroups of G is also a subgroup of G.

5. Let H be a subgroup of G. The left (right) cosets of H partition G and have cardinality $|H|$.

6. The *center* of G is the set $Z(G)$ of elements of G that commute with all of G, namely
$$Z(G) = \{z \in G \,|\, gz = zg \,\forall g \in G\}.$$
   a) Prove that $Z(G)$ is a subgroup of G.
   b) Show that $Z(G)$ is a normal subgroup of G.

7. Consider the map $\phi : G \to \mathrm{Aut}(G)$ defined by $\phi(g) = \varphi_g$, where φ_g is the automorphism of G defined by $\varphi_g(h) = ghg^{-1}$.
   a) Show that ϕ is a homomorphism.
   b) Show that $\ker(\phi) = Z(G)$.
   c) By the First Isomorphism Theorem, conclude that $G/Z(G) \cong \mathrm{Inn}(G)$.

3 Basic Group and Ring Properties

In this section, we focus on group actions and their role in determining general properties of order. We introduce rings and prove several easy results.

Definition 3.1. Let G be a group and S any nonempty set. A *left group action* is a homomorphism from G to the group of permutations of S; in particular, we regard the action as a map $\cdot : G \times S \to S$ such that $g \cdot (h \cdot s) = (gh) \cdot s$ and $e \cdot s = s$. The *orbit* of a point $s \in S$ is the set $Gs := \{g \cdot s \,|\, g \in G\}$, and the *stabilizer* of s is the set $G_s := \{g \in G \,|\, g \cdot s = s\}$.

For example, the cyclic group of order two, C_2, always acts by exchanging certain pairs of elements and fixing the remainder. Group actions often arise in geometry – as transformations of the plane; for instance, C_2 may act as reflection along a line, or 180° rotation about a point. A group acts on itself, naturally, via its multiplication rule. Another natural group action is given by *conjugation* [4].

It is easily checked that the stabilizer G_s is a subgroup of G. To emphasize this structure, we call G_s the *stabilizer subgroup* of s. As is the case with groups, the $\cdot$ is often omitted when its context is clear. As an introduction to the axiomatics of group actions, let us prove a basic result about actions.

Proposition 3.1. *The orbits of a group action $G \times S \to S$ partition S; the set of orbits is denoted S/G and called the* orbit space.

Proof. We check for reflexivity, symmetry, and transitivity in the relation $s \sim t$, where $s, t \in S$, that indicates whether $t \in Gs$. We have $s \sim s$ as $s = es \in Gs$. Now if $s \sim t$, then $t \in Gs$, so that there exists an element $g \in G$ such that $t = gs$. It follows that

$$s = es = (g^{-1}g)s = g^{-1}(gs) = g^{-1}t,$$

so that $s \in Gt$ or $t \sim s$ as well. Finally, if $s \sim t$ and $t \sim u$, then $t = gs$ and $u = ht$ for some $g, h \in G$, so that

$$u = ht = h(gs) = (hg)s.$$

It follows that $u \in Gs$, so that $s \sim u$, as required. □

Thus, a group action provides equivalence classes (its orbits) on its image. Two elements belonging to the same equivalence class are often said to be identified. Understanding structures is often key in mathematics. A group action $G \times S \to S$ can make S into a *graph*. Treating the set S as vertices and drawing edges between pairs of elements sharing an orbit, we naturally form a graph comprising a *disjoint union* of *cliques*. The following counting theorem is equally easy to prove, yet in its generality underlies many interesting results.

Theorem 3.2 (Orbit-Stabilizer Theorem)**.** *Suppose G acts on S. Fix a point $s \in S$ arbitrarily, and let L_s be the set of left cosets of G_s in G. Then the map $f : Gs \to L_s$ defined by $gs \mapsto gG_s$ is well defined and bijective. In particular,*

$$|Gs| = [G : G_s] \quad \textit{and} \quad |Gs| \cdot |G_s| = |G|$$

for all $s \in S$.

Proof. Suppose $gs = hs$. Then

$$s = (g^{-1}g)s = g^{-1}(gs) = g^{-1}(hs) = (g^{-1}h)s,$$

so that $g^{-1}h \in G_s$. It follows that $g^{-1}hG_s = G_s$, so that $hG_s = gG_s$, proving that the map is well defined. By its definition, the map is surjective. Injectivity

can be established by noting that if $gG_s = hG_s$, then $g^{-1}h \in G_s$, so that $(g^{-1}h)s = s$. Finally,

$$hs = (g(g^{-1}h))s = g((g^{-1}h)s) = gs,$$

as desired. Now observe that $|Gs| = |L_s| = [G : G_s]$. It is easy to see that the same number of elements of G carry s to each element of its orbit Gs, so that the equivalent formula $|Gs| \cdot |G_s| = |G|$ is obtained. □

As an application, one can easily deduce Burnside's lemma, a combinatorial lemma with a colorful history. For our present purposes, however, we deduce a famous theorem of Lagrange.

Corollary 3.3 (Lagrange's Theorem)**.** *If H is a subgroup of a finite group G, then $|G| = |H| \cdot [G : H]$. In particular, the order of the group is divisble by the order of any subgroup or any group element.*

Proof. Let L_H be the set of cosets of H. Then G acts on L_H by group multiplication. The stabilizer G_H is just H, and the orbit of H is all of H's cosets. By the Orbit-Stabilizer Theorem, we have

$$|G| = |G_H| \cdot |L_H| = |H| \cdot [G : H].$$

Observe that the index $[G : H]$ is a positive integer. To finish the proof, note that any group element by itself generates a cyclic subgroup of the same order. □

When the extra structure of a second operation is introduced, we move into the theory of rings.

Definition 3.2. A *ring* is a set R together with commutative and associative binary operations $+, \cdot : R \times R \to R$ such that addition makes R into an abelian group and multiplication distributes over addition. Elements having multiplicative inverses are called *units,* and the set of all units is denoted $R^\times$.

The set of integers $\mathbb{Z}$ is a ring, and so is the set of Gaussian integers $\mathbb{Z}[i]$. The latter example shows that formal elements can be added to a ring. The polynomials with integer coefficients over a single variable are denoted by $\mathbb{Z}[x]$; they, too, form a ring.

Often, multiplicative commutativity is dropped, but we maintain the axiom for the sake of simplicity. We remark also that the set of units $R^\times$ forms a group under multiplication in R; accordingly, $R^\times$ is often called the *group of units.*

Definition 3.3. An *ideal* is a subset $I \subset R$ that is closed under addition and satisfies $rI \subset I$ for all $r \in R$. An ideal P is called *prime* if whenever $r, s \in R$ and $rs \in P$ then either r or s lies in P. Many closed operations on ideals are defined: intersections, sums, and products of two ideals in the same ring are also ideals, where the latter two are defined as

$$I + J := \{i + j \,|\, i \in I, j \in J\} \text{ and}$$
$$IJ := \{i_1 j_1 + \cdots + i_n j_n \,|\, i_1, \ldots, i_n \in I; j_1, \ldots, j_n \in J\}.$$

Two ideals I and J in a ring R are called *coprime* if $I + J = R$. By analogy to cosets of normal subgroups, the additive cosets $a + I := \{a + i \,|\, i \in I\}$ yield the *quotient ring* R/I, where addition and multiplication are defined as $(a + I) + (b + I) = (a + b) + I$ and $(a + I)(b + I) = (ab) + I$.

In the integers, the multiples of 2 are denoted $2\mathbb{Z}$ and form an ideal. The ideal aR is often abbreviated (a); still more generally, the ideal $a_1R + \cdots + a_kR$ is written $(a_1, \ldots, a_k)$. Ideal addition respects this notation, in that $(a) + (b) = (a, b)$, as can be quickly verified. Multiplication of ideals obeys $IR = I$. More curiously, multiplication also obeys $IJ \subset I \cap J$.

Now let m and n be positive integers and let p be a prime. Because m divides n if and only if $n \in m\mathbb{Z}$, the definitions of prime and coprime ideals generalize the familiar arithmetic notions. If p divides the product mn, or equivalently $mn \in p\mathbb{Z}$, then p divides m or p divides n, so that $\{m, n\} \cap p\mathbb{Z} \neq \emptyset$. If m and n are coprime, then there exist integers s and t such that $ms + nt = 1$. It follows that $1 \in m\mathbb{Z} + n\mathbb{Z}$, so that $m\mathbb{Z} + n\mathbb{Z} = \mathbb{Z}$. Perhaps not surprisingly, $(2)(3) = (6)$ in the ring of integers. In many rings, ideal factorization generalizes our notion of elementary factorization. However, we might equally well write $(2) + (3) = (1)$.

Let us recall the fundamental idea of a group homomorphism; it extends in a natural way to rings.

Definition 3.4. A *ring homomorphism* is a map $\varphi : R \to S$ between two rings R and S such that $\varphi(1_R) = 1_S$ and, for all $a, b \in R$,

$$\varphi(a + b) = \varphi(a) + \varphi(b) \quad \text{and} \quad \varphi(ab) = \varphi(a)\varphi(b).$$

Again, if φ is bijective, then its inverse is also a ring homomorphism and the rings R and S are isomorphic.

By analogy with group theory, the *kernel* of a ring homomorphism $\varphi : R \to S$ is defined $\ker(\varphi) := \{r \in R \,|\, \varphi r = 0_S\}$. Likewise, the *image* of a ring homomorphism is defined $\mathrm{im}(\varphi) := \{s \in S \,|\, \exists r \in R : s = \varphi(r)\}$.

It is easy to check that the kernel is an ideal. Conversely, every ideal is the kernel of a homomorphism. For example, this can be seen from the natural quotient map $R \to R/I$ defined by $a \mapsto a+I$ for all $a \in R$. Thus, quotient groups G/H and quotient rings R/I come equipped with natural homomorphisms,

$$\overline{\phi} : G \to G/H \quad \text{and} \quad \overline{\varphi} : R \to R/I;$$

each identifies a group (ring) element with the (additive) coset containing the element.

We remark that there is a tendency to blur the distinction between a coset in the image and a representative (element) of the coset. When context is clear, this loses no information; however, we shall reserve $\overline{g}$ and $\overline{a}$ as typical quotient group and quotient ring elements (cosets) of which g and a are representatives, respectively.

After these introductory concepts, we can now understand "the integers modulo n" more precisely as the quotient ring $\mathbb{Z}/n\mathbb{Z}$.

Corollary 3.4 (Euler's Theorem)**.** *If a and n be coprime positive integers, then $n \mid a^{\phi(n)} - 1$, where ϕ is Euler's totient function.*

Proof. There are $\phi(n)$ integers b such that $0 \leq b < n$ and $(b, n) = 1$; these are in bijective correspondence with units in $(\mathbb{Z}/n\mathbb{Z})^\times$. For, given such a b, there exist integers s and t such that $bs + nt = 1$; passing to the quotient ring, we have

$$\overline{1} = \overline{bs + nt} = \overline{bs} + \overline{nt} = \overline{b}\overline{s} + \overline{n}\overline{t} = \overline{b}\overline{s},$$

so that $\overline{b}$ is a unit. Conversely, any integer not relatively prime to n cannot have a multiplcative inverse in $\mathbb{Z}/n\mathbb{Z}$. Now apply Lagrange's Theorem. □

The following is an immediate consequence.

Corollary 3.5 (Fermat's Little Theorem)**.** *If a is an integer and p is prime, then $p \mid a^p - a$.*

Proof. Either $p \mid a$ or $p \nmid a$. In the latter case, note that $\phi(p) = p - 1$ and use Euler's Theorem. □

Nontrivial information is conveyed by homomorphisms. For example, it is sometimes easier to study factorization in a quotient ring. Consider the quotient map

$$\pi : \mathbb{Z}[x] \to \mathbb{Z}[x]/p\mathbb{Z}[x] \cong (\mathbb{Z}/p\mathbb{Z})[x],$$

where p is a prime. Then $P(x) = Q(x)R(x) \Longrightarrow \overline{P(x)} = \overline{Q(x)}\,\overline{R(x)}$. With this in mind, a simple contradiction argument gives the following result.

Corollary 3.6 (Eisenstein's Criterion)**.** *If $P(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0$ and p is a prime such that $p \nmid a_n, p^2 \nmid a_0$, and $p \mid a_i$ for all $i = 0, \ldots, n-1$, then $P(x)$ is irreducible.*

Proof. For, $\overline{P(x)} = \overline{a_n}x^n$. If $P(x) = Q(x)\,R(x)$ where both Q and R are nonconstant, then both $\overline{Q(x)}$ and $\overline{R(x)}$ are divisible by x. Thus, $p \mid Q(0)$ and $p \mid R(0)$, contrary to $p^2 \nmid a_0$. □

The familiar division algorithm for polynomials in a single variable depends only on the existence of multiplicative inverses. Although the division algorithm applies to arbitrary fields, many ring elements do not have inverses, so the algorithm does not apply to division by any arbitrary polynomial in $R[x]$. However, it is easy to see that division by a polynomial with a unit as leading coefficient proceeds as usual.

Proposition 3.7 (Factor Theorem)**.** *Let $p(x) \in R[x]$ be a polynomial of degree n. For each element $a \in R$, there exists a polynomial $q(x)$ of degree $n - 1$ such that $p(x) - p(a) = (x - a)q(x)$.*

Proof. The division algorithm gives $p(x) - p(a) = (x - a)q(x) + r$, where $r \in R$ is constant. Plugging in $x = a$ shows that $r = 0$. □

Corollary 3.8 (Wilson's Theorem)**.** *For a prime p, we have $p \mid (p-1)! + 1$.*

Proof. Assume $p > 2$. Consider the polynomial $p(x) = x^{p-1} - 1$ in the field $\mathbb{Z}/p\mathbb{Z}$. By Fermat's Little Theorem, the roots of $p(x)$ are $1, 2, \ldots, p-1$. Using the factor theorem to expand, we have

$$p(x) = (x-1)(x-2)\cdots(x-(p-1)) = x^{p-1} - \cdots + 1 \cdot 2 \cdots (p-1).$$

The result follows by comparing this expansion to $x^{p-1} - 1$. □

Exercises

Let G be a finite group.

1. Show that *conjugation* gives a group action $G \times G \to G$ defined by $(g, h) \mapsto h^{-1}gh$.

2. Let $G \times S \to S$ be a group action. Show that for any $s \in S$ the stabilizer $G_s := \{g \mid gs = s\}$ is a subgroup of G.

3. (Burnside's lemma) Let $G \times S \to S$ be a group action and let S^g denote the subset of S fixed by $g \in G$. Prove that the number of orbits is given by
$$|S/G| = \frac{1}{|G|} \sum_{g \in G} |S^g|.$$

4. Determine the number of ways to color the vertices of a cube with k-colors, where two colorings are the same if one can be rotated into the other.

   a) Argue that the rotations of a cube have a natural composition law making them into a group.

   b) Classify each rotations as one of several types.

   c) Deter the colorings fixed by each type of rotation and apply Burnside's lemma.

5. Determine the prime ideals in $\mathbb{Z}[x]$ and $\mathbb{R}[x]$.

6. (Gauss's Lemma) A *primitive polynomial* over $\mathbb{Z}$ is a polynomial $p(x) \in \mathbb{Z}[x]$ such that the coefficients of $p(x)$ have no common divisors other than the units ± 1.

   a) Show that the product of two primitive polynomials is a primitive polynomial. *Hint*: examine the converse in a quotient ring.

   b) Let $f(x) \in \mathbb{Z}[x]$ be a primitive polynomial that factors as $f(x) = g(x)h(x)$ for some nonconstant polynomials $g(x), h(x) \in \mathbb{Q}[x]$. Argue that without loss of generality, there exist positive integers m and n such that $mg(x)$ and $nh(x)$ are primitive polynomials in $\mathbb{Z}[x]$.

   c) Deduce that mn is a unit.

   d) Conclude that a polynomial in $\mathbb{Z}[x]$ factors in $\mathbb{Z}[x]$ if and only if it factors in $\mathbb{Q}[x]$. Note that a similar result where $\mathbb{Z}$ is replaced by a *unique factorization domain* R and $\mathbb{Q}$ is replaced by the field of fractions F of R.

7. Let a and n be coprime positive integers. Prove Euler's theorem, that n divides $a^{\phi(n)} - 1$, by examining multiplication by a in the group of units $(\mathbb{Z}/n\mathbb{Z})^\times$.

8. Let p be a prime. Prove Fermat's little theorem by considering an action of C_p, the cyclic group of order p, on the Cartesian product $\{1, 2, 3, \ldots, a\}^p$.

9. Let p be a prime. Show that the *cyclotomic polynomial*

$$\phi_p(z) = z^{p-1} + \cdots + z + 1$$

is irreducible over $\mathbb{Z}$. *Hint:* consider transformations that preserve irreducibility.

10. (First Sylow Theorem) Let p be a prime, let n be a positive integer, and suppose the order of G is an integer divisible by p^n.

a) Consider the collection $\mathcal{C} := \{S \subset G : |S| = p^n\}$, let $|G| = p^n r$, and suppose that p^m is the largest power of p dividing r. Show that p^m is the largest power of p that divides $|\mathcal{C}|$.

b) Consider the group action $G \times \mathcal{C} \to \mathcal{C}$ defined by left multiplication, $S \mapsto gS = \{gh : h \in S \subset G\}$ for each $g \in G$. Show that for some set $S \in \mathcal{C}$, the orbit GS has cardinality not divisible by p^{m+1}.

c) Using the Orbit-Stabilizer Theorem, deduce that the stabilizer of this subset is large, namely that $p^n \leq |G_S|$.

d) By considering an element $s \in S$, argue that $|S| \geq |G_S|$.

e) Conclude that G has a subgroup of order p^n.

11. (Second Sylow Theorem) A *p-group* is finite group whose order is a power of p. A *Sylow p-subgroup* of G is a subgroup of G is a p-group whose order p^k is the largest power of p dividing $|G|$.

a) Let a p-group H act on a finite set S and define the subset $S_0 := \{s \in S \,|\, H_s = H\}$ of elements fixed by H. Use the Orbit-Stabilizer Theorem to show that $|S| \equiv |S_0| \pmod{p}$.

b) Let P be a Sylow p-subgroup of G and let H be p-subgroup of G. Consider the collection $\mathcal{C}$ of left cosets of P and let H act on $\mathcal{C}$ by left multiplication. Deduce that some coset gP is fixed by each element of H.

c) Conclude that $g^{-1}Hg \subset P$.

d) Argue that the Sylow p-subgroups are conjugate to one another and thus isomorphic.

12. (Third Sylow Theorem) Let n_p denote the number of Sylow p-subgroups of a finite group G, and let q be their common order. Show that n_p divides $|G|/q$ and that $n_p \equiv 1 \pmod{p}$.

4 Structure Theorems

Many of the theorems in the previous section have in common that they give general properties of specific group elements. This section considers more detailed structures and addresses a dual question, whether there exist group elements with specific properties. With the goal of constructing group elements of particular orders, we shall proceed combinatorially.

For the remainder of this paper, G denotes a group, and m and n denote positive integers. We will abbreviate $\mathbb{Z}/n\mathbb{Z}$ by $\mathbb{Z}/n\mathbb{Z}$ and $(\mathbb{Z}/n\mathbb{Z})^\times$ by $(\mathbb{Z}/n\mathbb{Z})^\times$. The order of an element $g \in G$ is written $O_G(g)$, or simply $O(g)$ where the context is clear. Recall that the order of an element is the least positive integer such that $g^{O(g)} = 1$, or ∞ if no such integer exists. For any group element $g \in G$, the cyclic subgroup generated by g is defined by $\langle g \rangle := \{\dots, g^{-1}, 1, g, g^2, \dots\}$; if $O(g) < \infty$, then $\langle g \rangle = \{1, g, \dots, g^{O(g)-1}\}$. From the definition of $O(g)$, we have $g^n = 1$ if and only if $O(g) \mid n$. In particular, $\{n \in \mathbb{Z} \mid g^n = 1\} = O(g)\mathbb{Z}$.

Proposition 4.1. *Let a group element $g \in G$ have finite order. Then the order of each element of the subgroup $\langle g \rangle$ divides $O(g)$, and every divisor of $O(g)$ is the order of an element of $\langle g \rangle$.*

Proof. Consider $a \in \langle g \rangle$; there exists a unique integer $0 \leq m < O(g)$ such that $a = g^m$. Then $a^n = (g^m)^n = g^{mn}$, so that $a^n = 1 \iff O(g) \mid mn$. Writing $n = O(g)$, it follows that $a^{O(g)} = 1$, and $O(a) \mid O(g)$. Moreover, $O(a) = \operatorname{lcm}(O(g), m)/m$. Writing $O(g) = d_1 d_2$ and putting $m = d_2$ yields $O(a) = \operatorname{lcm}(O(g), d_2)/d_2 = O(g)/d_2 = d_1$, as desired. □

The proposition above shows that we can easily find group elements with small orders. In particular, the reader familiar with the *poset* of divisibility relations in the positive integers might notice that we have just shown that the set of group element orders is an *order ideal* [5]. A more difficult investigation concerns the existence of elements with large orders, a matter to which we now turn.

Proposition 4.2. *If G is abelian and contains elements of coprime orders, m and n, then G contains an element of order mn.*

Proof. Let $a, b \in G$ be such that $O(a) = m$ and $O(b) = n$; we claim that the product ab is satisfactory. Observe that $O(ab) \mid mn$, since

$$(ab)^{mn} = a^{mn} b^{mn} = (a^m)^n (b^n)^m = 1.$$

On the other hand, $a^{O(ab)} = b^{-O(ab)} \in \langle a \rangle \cap \langle b \rangle$. Because 1 is the only common positive divisor of m and n, the orders of $a^{O(ab)}$ and $b^{-O(ab)}$ must be 1; that is, both elements are the identity. Then $O(ab)$ is divisible by $O(a)$ and $O(b)$, as required. □

We should like to find elements of coprime order. Knowing that order lowering is an easy process, we might hope for the following compromise.

Proposition 4.3. *If G is abelian and contains elements of order m and n, then G contains an element of order $\text{lcm}(m,n)$.*

Proof. Consider k so large that the index set $I := \{1,\dots,k\}$ defines exponents such that $m =: \prod_{i\in I} p_i^{a_i}$ and $n =: \prod_{i\in I} p_i^{b_i}$, where the sequence $p_1, p_2, p_3, \dots$ is the primes in increasing order. Define $I_A := \{i \in I \mid a_i \geq b_i\}$, the set of indices having maximal exponents in the factorization of m, and denote the complement $I_B := I \setminus I_A$. By Proposition 3.1, G contains elements of orders $\overline{m} := \prod_{i\in I_A} p_i^{a_i}$ and $\overline{n} := \prod_{i\in I_B} p_i^{b_i}$. Because $I_A \cap I_B = \emptyset$, the orders $\overline{m}$ and $\overline{n}$ are coprime, so that Proposition 3.2 gives an element of G having order

$$\overline{m}\,\overline{n} = \prod_{i\in I_A} p_i^{a_i} \prod_{i\in I_B} p_i^{b_i} = \prod_{i\in I} p_i^{\max\{a_i,b_i\}} = \text{lcm}(m,n),$$

as desired. □

Motivated by our ability to usefully combine any pair of elements with finite order in an abelian group, we introduce the following definitions that amalgamate all of the elements.

Definition 4.1. A positive integer m is a *universal order* for the group G if $g^m = 1$ for all $g \in G$. The smallest such integer is the *least universal order* of G and is denoted O_G.

A group need not have any universal order. Consider, for example, the group of integers under addition. In fact, there exist groups in which every element has finite order but no universal order exists; consider $\mathbb{Q}/\mathbb{Z}$ under addition, for example.

But not all is lost. Because the positive integers are well-ordered, if a group has a universal order then it has a least universal order. And there exist conditions that guarantee a universal order.

Proposition 4.4. *Every finite abelian group G contains an element of order O_G, and O_G divides $|G|$.*

Proof. Consider group elements $g_1,\dots,g_k \in G$. By Lagrange's Theorem, these elements have finite orders $O(g_1),\dots,O(g_k)$ dividing $|G|$. It follows that the least common multiple of these elements is the least universal order. In particular, it divides $|G|$. Now observe that

$$\text{lcm}(O(g_1),O(g_2),O(g_3)) = \text{lcm}(\text{lcm}(O(g_1),O(g_2)),O(g_3)).$$

With this observation, we construct group elements $h_1,\dots,h_k \in G$ such that the orders $O(h_i) = \text{lcm}(O(g_1),\dots,O(g_i))$ by applying Proposition 3.3 to the pairs $\{h_i, g_{i+1}\}$. Then h_k has least universal order, as required. □

Definition 4.2. Let R be a ring. An nonzero element $a \in R$ is called a *zerodivisor* (pl. zerodivisors) if there exists another nonzero element $b \in R$ such that $ab = 0$. A ring with no zerodivisors is an *integral domain.*

For example, $\mathbb{Z}/3\mathbb{Z}$ is an integral domain but $\mathbb{Z}/4\mathbb{Z}$ is not (since $\overline{2}^2 = \overline{0}$). In fact, any field is an integral domain, because nonzero elements have multiplicative inverses. For, if $1 = a^{-1}a$ and $ab = 0$, then

$$b = (a^{-1}a)b = a^{-1}(ab) = a^{-1}0 = 0.$$

This is the crucial property behind our next statement.

Lemma 4.5 (Root Theorem)**.** *Let R be an integral domain. A polynomial $p(x) \in R[x]$ of degree n has at most n roots, including multiplicity.*

Proof. Use induction. The case $n = 1$ is trivial. So let $n > 1$ and suppose $p(x)$ has a root x_0. Then the division algorithm gives $p(x) = (x - x_0)q(x) + r$ where $q(x)$ has degree $n - 1$ and r is constant. Plugging in $x = x_0$ gives $r = 0$. Now if a is a root of $p(x)$, then $(a - x_0)q(a) = 0$, implying that $a = x_0$ or a is a root of $q(x)$. It follows that $p(x)$ has at most one more root than $q(x)$, as required. □

We are at last ready to discuss modular arithmetic.

Proposition 4.6. *$(\mathbb{Z}/p\mathbb{Z})^\times$ is a cyclic group.*

Proof. Introduce the least universal order $m := O_{(\mathbb{Z}/p\mathbb{Z})^\times}$, and consider the polynomial $p(x) := x^m - 1$. On the one hand, Proposition 4.4 shows that $m \,|\, p-1$; hence, $m \leq p-1$. On the other hand, by definition, all $p-1$ elements of $(\mathbb{Z}/p\mathbb{Z})^\times$ are roots of $p(x)$. The Root Theorem shows that $m \geq p - 1$. Thus, $m = p - 1$. And it follows that $(\mathbb{Z}/p\mathbb{Z})^\times$ contains an element of order $p - 1$, which is the generator we sought. □

One might ask whether the primality condition can be relaxed in the above proposition. In fact, precisely the same proof holds for all finite fields, which necessarily have prime power orders [4]. However, the finite field $\mathbb{F}_{p^k}$ and the quotient ring $\mathbb{Z}/p^k\mathbb{Z}$ cease being isomorphic for $k > 1$ (for instance, p becomes a zerodivisor in the ring), so more machinery is needed to assess the usual modular arithmetic.

Introducing powers of p leads to consideration of $z^p - 1$; curiously, it is easier to consider compare such a difference to zero than equate z^p with 1 in its own right. Naturally, one has the factorization $z^p - 1 = (z - 1)(z^{p-1} + \cdots + z + 1)$, to which we turn.

Lemma 4.7. *Let p an odd prime. Then for each integer x, the cyclotomic polynomial $\phi_p(x) = x^{p-1} + \cdots + x + 1$ is not divisible by p^2.*

Proof. Suppose for contradiction's sake that $p^2 | \phi_p(a)$ for some integer a. Then in modulo p,

$$\begin{aligned} 0 &\equiv (a-1)(a^{p-1} + \cdots + a + 1) \\ &= a^p - 1 \\ &\equiv a - 1. \end{aligned}$$

Thus, $a = 1 + kp$ for some integer k. But then in modulo p^2,

$$\begin{aligned} 1 + a + \cdots + a^{p-1} &= 1 + (1 + kp) + \cdots + (1 + kp)^{p-1} \\ &\equiv 1 + (1 + kp) + \cdots + (1 + (p-1)kp) \\ &= p(1 + kT_{p-1}), \end{aligned}$$

where $T_{p-1} = 1 + \cdots + (p-1) = (p-1)p/2$. Since p is odd, p divides T_p, so that $\phi_p(a) \equiv p \pmod{p^2}$, a contradiction.

□

The following gem is immediate.

Lemma 4.8 (Lifting Lemma)**.** *Let p be an odd prime. If a and k are nonnegative integers such that $p^k \nmid a - 1$, then $p^{k+1} \nmid a^p - 1$.*

Proof. One writes $a^p - 1 = (a-1)(a^{p-1} + \cdots + a + 1)$, and observes that p divides the second factor at most once. □

Returning, we need a means of comparing different moduli. Observe that for all pairs of positive integers m and n there exists a natural ring homomorphism $\varphi_{m,n} : \mathbb{Z}/mn\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$ due to the ideal containment $mn\mathbb{Z} \subset n\mathbb{Z}$. In particular, each coset in $\mathbb{Z}/n\mathbb{Z}$ is partitioned by additive cosets in $\mathbb{Z}/mn\mathbb{Z}$; these partitions comprise the preimages. For example, consider the map $\varphi_{3,2} : \mathbb{Z}/6\mathbb{Z} \to \mathbb{Z}/2\mathbb{Z}$. The preimage of $\overline{1} = 1 + 2\mathbb{Z}$ is the collection $\varphi_{3,2}^{-1}(\overline{1}) = \{1 + 6\mathbb{Z}, 3 + 6\mathbb{Z}, 5 + 6\mathbb{Z}\}$. Because $1 \in \mathbb{Z}/mn\mathbb{Z} \mapsto \overline{1} \in \mathbb{Z}_n$, it follows that $O_{(\mathbb{Z}/n\mathbb{Z})^\times}(\overline{a}) \,|\, O_{(\mathbb{Z}/mn\mathbb{Z})^\times}(a)$ for all $a \in (\mathbb{Z}/mn\mathbb{Z})^\times$. That is, the multiplicative order of a unit in modulo mn is divisible by the order it assumes in modulo n.

As it turns out, $\left(\mathbb{Z}/p^k\mathbb{Z}\right)^\times$ is almost always cyclic.

Proposition 4.9. *The group $\left(\mathbb{Z}/p^k\mathbb{Z}\right)^\times$ is cyclic if and only if p is odd or $k \leq 2$.*

Proof. Consider the case $p = 2$. Observe that both $(\mathbb{Z}/2\mathbb{Z})^\times$ and $(\mathbb{Z}/4\mathbb{Z})^\times$ are cyclic while $(\mathbb{Z}/8\mathbb{Z})^\times$ is not; we have $1^2 \equiv 3^3 \equiv 5^2 \equiv 7^2 \equiv 1 \pmod 8$. Now take an integer $k > 3$ and consider the natural homomorphism $\left(\mathbb{Z}/2^k\mathbb{Z}\right)^\times \to (\mathbb{Z}/8\mathbb{Z})^\times$. Because the map is surjective, a generator of $\left(\mathbb{Z}/2^k\mathbb{Z}\right)^\times$ would map to a generator in $(\mathbb{Z}/8\mathbb{Z})^\times$, of which there are none.

Now assume $p > 2$ and consider $k = 2$. Let $g \in \left(\mathbb{Z}/p^2\mathbb{Z}\right)^\times$ be such that $\varphi_{p,p}(g)$ generates $(\mathbb{Z}/p\mathbb{Z})^\times$. Then $O_{(\mathbb{Z}/p\mathbb{Z})^\times}(g) = p - 1$ and we have that $p - 1$ divides $O_{(\mathbb{Z}/p^2\mathbb{Z})^\times}(g + mp)$ for all m. By Lagrange's Theorem, the orders $\{O_{(\mathbb{Z}/p^2\mathbb{Z})^\times}(g + mp)\}_m$ are divisors of $|\left(\mathbb{Z}/p^2\mathbb{Z}\right)^\times| = \phi(p^2) = p(p-1)$. Thus, $O_{(\mathbb{Z}/p^2\mathbb{Z})^\times}(g + mp) \in \{p - 1, p(p-1)\}$ for all m. Suppose that g has order $p - 1$ in $\left(\mathbb{Z}/p^2\mathbb{Z}\right)^\times$. Then

$$(g + p)^{p-1} \equiv g^{p-1} + g^{p-2}p(p-1) \equiv 1 - pg^{p-2} \pmod{p^2}.$$

Because the image of g generates $(\mathbb{Z}/p\mathbb{Z})^\times$, g is coprime to p, and it follows that the order of $g+p$ is not $p-1$. Thus, at least one of g and $g+p$ is a generator of $\left(\mathbb{Z}/p^2\mathbb{Z}\right)^\times$.

To complete the proof for $k > 2$, we claim that an integer g that generates $\left(\mathbb{Z}/p^{k-1}\mathbb{Z}\right)^\times$ also generates $\left(\mathbb{Z}/p^k\mathbb{Z}\right)^\times$. By the natural homomorphism

$$\left(\mathbb{Z}/p^k\mathbb{Z}\right)^\times \to \left(\mathbb{Z}/p^{k-1}\mathbb{Z}\right)^\times,$$

we have that the order $p^{k-2}(p-1)$ of g in $\left(\mathbb{Z}/p^{k-1}\mathbb{Z}\right)^\times$ divides the order of g in $\left(\mathbb{Z}/p^k\mathbb{Z}\right)^\times$. By Lagrange's theorem, $O_{(\mathbb{Z}/p^k\mathbb{Z})^\times}(g)$ divides $p^{k-1}(p-1)$. Finally, as $p^{k-1} \nmid g^{p^{k-3}(p-1)} - 1$, the lifting lemma gives $p^k \nmid g^{p^{k-2}(p-1)} - 1$. □

The strategy employed in final paragraph of the preceeding proof merits particular attention. We realize a strong and abstract deduction by essentially squeezing an equality case out of an inequality; observe how Lagrange's theorem acts as an upper bound while the natural homomorphism and lifting lemma provide a lower bound that happens to be sharp in the integer divisibility lattice.

All that remains in extending our primitive roots proposition is generalization from prime powers to arbitrary integers. The techniques to this end are varied and eclectic, and the best motivation comes from practice. We shall use the Chinese Remainder Theorem. However, to understand it and its usage, we require several notions from our earlier encounter with ring theory, which we recall presently. Two ideals $I, J \subset R$ are coprime if their sum is the entire overlying ring, or $I + J = R$. Ideal products are defined and obey the containment $IJ \subset I \cap J$. As might be hoped, addition and multiplication of ideals is commutative and distributive in a commutative ring. Quotient rings R/I are defined in terms of additive cosets, $r+I$, having operations $(a+I)+(b+I) = (a+b)+I$ and $(a+I)(b+I) = (ab)+I$ inherited from R.

Lastly, the direct product of the rings $R_1, \ldots, R_k$ is just the ring structure on the cartesian product obtained by allowing the factors to act coordinatewise; more precisely, the elements are k-tuples such as $(a_1, \ldots, a_k)$ and $(b_1, \ldots, b_k)$ where $a_i, b_i \in R_i$, and the operations are defined as

$$\begin{aligned}(a_1, \ldots, a_k) + (b_1, \ldots, b_k) &= (a_1 + b_1, \ldots, a_k + b_k),\\ (a_1, \ldots, a_k)(b_1, \ldots, b_k) &= (a_1 b_1, \ldots, a_k b_k).\end{aligned}$$

Note that the additive and multiplicative identities are $(0, \ldots, 0)$ and $(1, \ldots, 1)$.

Intuitively, the behavior of a ring element as it appears in multiple quotient rings should provide information about that element in some combined quotient ring, and vice versa. This is the content of the theorem we are about to prove.

Theorem 4.10 (Chinese Remainder Theorem)**.** *If R is a ring and $I_1, I_2, \ldots, I_k$ are pairwise coprime ideals, then $I_1 I_2 \cdots I_k = I_1 \cap I_2 \cap \cdots \cap I_k =: I$ and the map*

$$f : R/I \to R/I_1 \times \cdots \times R/I_k$$

defined by $f(r+I) := (r+I_1, \ldots, r+I_k)$ is a ring isomorphism.

Proof. First, we check that I is well defined. Observe that

$$R = \prod_{i \neq 1}(I_1 + I_i) \subset I_1 + \prod_{i \neq 1} I_i \subset R,$$

implying that $I_1 + I_2 I_3 \cdots I_n = R$. Now if J and K are coprime ideals, $J \cap K = (J \cap K)(J + K) \subset KJ + JK \subset JK$. But the reverse containment always holds, so $J \cap K = JK$. Using $J = I_1$ and $K = I_2$ as a base case together with the inductive step

$$I_1 \cap \left(\bigcap_{i=2}^{k} I_i \right) = I_1 \cap \left(\prod_{i=2}^{k} I_i \right) = \prod_{i=1}^{k} I_i,$$

we see that $I_1 I_2 \cdots I_n = I_1 \cap \cdots \cap I_n$. Hence, I is indeed well-defined. Now consider the ring homomorphism $\varphi : R \mapsto \prod_{i=1}^{n} R/I_i$ defined by

$$r \overset{\varphi}{\mapsto} (r + I_1, r + I_2, \ldots, r + I_n).$$

Note that $\ker(\varphi) = I_1 \cap \cdots \cap I_n = I$. To see that φ is surjective, we will find $r \in R$ such that $r \in 1 + I_1$ and $r \in \prod_{i \neq 1} I_i$; this is sufficient because, for such r, we have $\varphi(r) = (\overline{1}, \overline{0}, \ldots, \overline{0})$. But we already saw that $1 \in R = I_1 + I_2 \cdots I_n$, so $1 = j + j^*$ for some $j \in I_1, j^* \in I_2 \cdots I_n$ and $r = 1 - j = j^*$ has the properties we sought. Finally, the First Isomorphism Theorem [7] gives the central isomorphism in

$$R/I \cong R/\ker(\varphi) \cong \mathrm{im}(\varphi) = \prod_{i=1}^{n} R/I_i,$$

as desired. □

This abstract formulation has a more elementary corollary, often given the same name, in the integers.

Corollary 4.11 (Chinese Remainder Theorem)**.** *Let k be a positive integer and suppose that integers $m_1, \ldots, m_k$ are pairwise coprime. Then for each k-tuple $(n_1, \ldots, n_k)$ of integers such that $0 \leq n_i < m_i$ for each i there exists a unique integer n such that $0 \leq n < m_1 \cdots m_k$ and for all indices i,*

$$n_i \equiv n \pmod{m_i}.$$

Moreover, this correspondence is bijective.

Proof. Observe that if $i \neq j$ then there exist integers a and b such that $am_i + bm_j = 1$. Thus, $1 \in m_i\mathbb{Z} + m_j\mathbb{Z}$, implying that $m_i\mathbb{Z} + m_j\mathbb{Z} = \mathbb{Z}$. Moreover, $(m_i\mathbb{Z})(m_j\mathbb{Z}) = (m_i m_j)\mathbb{Z}$. The ideals $\{m_i\mathbb{Z}\}_i$ meet the criteria in the Theorem 3.1 and this in turn gives the ring isomorphism

$$\mathbb{Z}/m_1 \cdots m_k\mathbb{Z} \cong \mathbb{Z}/m_1\mathbb{Z} \times \cdots \times \mathbb{Z}/m_k\mathbb{Z}.$$

Identifying each additive coset $a + b\mathbb{Z}$ with the unique nonnegative integer less than $|b|$ it contains, we reduce this bijection to the claims. □

There is a subtlety involved in using the Chinese Remainder Theorem to describe $(\mathbb{Z}/n\mathbb{Z})^\times$. The ring isomorphism $\mathbb{Z}/m_1 \cdots m_k\mathbb{Z} \cong \mathbb{Z}/m_1\mathbb{Z} \times \cdots \times \mathbb{Z}/m_k\mathbb{Z}$ respects both additive and multiplicative structures, but what we actually desire is a group isomorphism on the respective groups of units. Note that a ring isomorphism $\phi : R \to S$ carries units to units, since $\phi(a)\phi(a^{-1}) = \phi(aa^{-1}) = \phi(1_R) = 1_S$. Thus, dropping the additive property of ϕ and restricting the domain and image, ring isomorphisms induce isomorphisms between unit groups. Finally, the unit group of a direct product is the direct product of the corresponding unit groups, so that

$$(\mathbb{Z}/n\mathbb{Z})^\times \cong \left(\prod_{i=1}^{k} \mathbb{Z}/m_i\mathbb{Z}\right)^\times \cong \prod_{i=1}^{k} (\mathbb{Z}/m_i\mathbb{Z})^\times .$$

We are at last ready for this section's main result.

Theorem 4.12 (Primitive Root Theorem)**.** *The group $(\mathbb{Z}/n\mathbb{Z})^\times$ is cyclic if and only if $n \in \{2, 4\}$ or $n \in \{p^k, 2p^k\}$ for some positive integer k and odd prime p.*

Proof. Write $n = \prod_{i=1}^{m} p_i^{a_i}$ and consider a residue class $g \in (\mathbb{Z}/n\mathbb{Z})^\times$. By the Chinese Remainder Theorem,

$$O_{(\mathbb{Z}/n\mathbb{Z})^\times}(g) = \operatorname{lcm}(O_{\left(\mathbb{Z}/p_1^{a_1}\mathbb{Z}\right)^\times}(g), \ldots, O_{(\mathbb{Z}/p_m^{a_m}\mathbb{Z})^\times}(g)).$$

If g is a generator in modulo n, then g is a generator in modulo $p_i^{a_i}$ for each i. Under this hypothesis, the equation above reduces to

$$\begin{aligned} \prod_{i=1}^{m}(p_i - 1)p_i^{a_i - 1} &= \phi(n) \\ &= O_{(\mathbb{Z}/n\mathbb{Z})^\times} \\ &= \operatorname{lcm}(\{O_{\left(\mathbb{Z}/p_i^{a_i}\mathbb{Z}\right)^\times}(g)\}_i) \\ &= \operatorname{lcm}(\{(p_i - 1)p_i^{a_i - 1}\}_i), \end{aligned}$$

so that the numbers $\{(p_i - 1)p_i^{a_i - 1}\}_i$ are pairwise coprime. Thus, if $(\mathbb{Z}/n\mathbb{Z})^\times$ is cyclic, then n cannot be divisible by both 4 and an odd prime or divisible by two odd primes. Of the remaining candidates, only those of the form $2p^k$ are new. But $\phi(2p^k) = \phi(p^k)$, so that any odd integer generating modulo p^k also generates modulo $2p^k$. □

Now that we have answered which moduli have generators, we turn to a dual question that merits attention: given an integer, for which moduli is it a generator? Once more, it is natural to study the query for prime moduli. However, progress in this direction has proven very difficult. The following conjecture was made by Emil Artin in 1927 [8].

Conjecture 4.1 (Artin's Conjecture)**.** *Suppose $a \neq -1$ is a nonsquare integer. Let P denote the set of primes, and denote by $S(a)$ the set of primes p for which*

$\overline{a}$ generates $(\mathbb{Z}/p\mathbb{Z})^\times$. Then

1. *$S(a)$ has positive density inside the set of primes, in particular $|S(a)| = \infty$;*
2. *if a are squarefree, then the density equals the* ***Artin constant***

$$C_{Artin} = \prod_{p \in P} \left(1 - \frac{1}{p(p-1)}\right) = 0.3739558136\ldots.$$

Artin's conjecture has resisted a complete solution for over eighty years. Nonetheless, several notable achievements have been realized. In 1967, Hooley gave conditional proof establishing the conjecture assuming parts of the Generalized Riemann Hypothesis [9]. A 1984 paper [10] by Gupta and Murty showed that the conjecture holds for infinitely many a. In 1986, Heath-Brown extended this infinitude of solutions, showing, amazingly, that there are at most two prime numbers a violating the claims [11]. It is not known, however, which value(s), if any, fail. So, while Artin's conjecture is known to hold for at least one value $a \in \{3, 5, 7\}$, we still cannot tell which, and in fact, no particular value of a has yet been proven to satisfy the conditions. †

Exercises

1. (USAMO 2005) Prove that the system

$$\begin{aligned} x^6 + x^3 + x^3y + y &= 147^{157} \\ x^3 + x^3 + y^2 + y + z^9 &= 157^{147} \end{aligned}$$

   has no solutions in integers x, y, and z.

2. (IMO Short List 1997/N4) Show that if an infinite arithmetic progression of positive integers contains a square and a cube, it must contain a sixth power.

3. (Euler's Criterion) The *Legendre symbol* is defined as follows: if p is an odd prime and a is an integer,

$$\left(\frac{a}{p}\right) = \begin{cases} 0 & \text{if } p \text{ divides } a \\ 1 & \text{if } a \text{ is a square in } \mathbb{Z}/p\mathbb{Z} \\ -1 & \text{otherwise.} \end{cases}$$

   Show that

$$\left(\frac{a}{p}\right) \equiv a^{(p-1)/2} \pmod{p}.$$

4. Let p be an odd prime. Show that for any integers a, b,

$$\left(\frac{a}{p}\right)\left(\frac{b}{p}\right) = \left(\frac{ab}{p}\right).$$

   Thus, the Legendre symbol is multiplicative in its top argument.

5. Recall that a *Fermat prime* is a prime number of the form $2^{2^n}+1$. An odd prime p is called *orderly* if for each integer a,

$$\left(\frac{a}{p}\right) = -1 \iff a \text{ generates } \mathbb{Z}/p\mathbb{Z}.$$

Prove that an odd prime is orderly if and only if it is a Fermat prime.

6. (IMO Short List 1998) Find all positive integers n such that $2^n - 1$ divides $m^2 + 9$ for some integer m.

7. (Italian TST 2006/5) For each positive integer n define the set $A_n = \{a \in \mathbb{Z} : 1 \leq a \leq n; n \,|\, a^n + 1\}$.

   a) Find all n such that $A_n \neq \emptyset$.

   b) Find all n such that $|A_n|$ is even and nonzero.

   c) Does there exist an integer n such that $|A_n| = 130$?

8. Let p be a prime. Denote by S the set of all primitive roots in $\mathbb{Z}/p\mathbb{Z}$. Compute $\sum_{a \in S} a \pmod p$. *Hint:* investigate a connection with the Möbius function and cyclotomic polynomials.

9. (?) Determine whether or not there exist arbitrarily long sequences of consecutive positive integers no two of which have the same number of prime divisors.

5 Nullstellensatz

The following discussion is heavily inspired by, indeed contains many excerpts from, a paper by Noga Alon [12]. By now, the reader should be capable of following the excellent presentation there and is therefore encouraged to read that literature.

The theory in this section generalizes familiar results about zeros of univariate polynomials, which we recall presently. Consider a polynomial f having coefficients in an integral domain R. If f can be written as $a_n x^n + \cdots + a_1 x + a_o$ and there are $n+1$ values x such that $f(x) = 0$, then $f = 0$ identically. If $r_1, \ldots, r_k$ are roots of the polynomial then $f(x) = h(x)(x - r_1) \cdots (x - r_k)$ for some polynomial $h(x)$ with coefficients in R. Moreover, the degree of h will be $n - k$. Finally, if f has degree n, any set of $n+1$ values in its coefficient ring contains a non-zero.

Lemma 5.1 (Brick Lemma)**.** *Fix a polynomial $P \in R[x_1, \ldots, x_n]$, where R is an integral domain. Suppose that the degree of P as a polynomial in x_i is at most t_i for $1 \leq i \leq n$, and let $S_i \subset R$ be such that $|S_i| \geq t_i + 1$. If $P(s_1, \ldots, s_n) = 0$ for all tuples $(s_1, \ldots, s_n) \in S_1 \times \cdots \times S_n$, then $P = 0$.*

The proof of this lemma is left as an exercise. This lemma is simply the natural generalization of the familar univariate result and expresses a property of the zero polynomial.

Theorem 5.1 (Ideal Decomposition)**.** *Fix a polynomial $f \in R[x_1, \ldots, x_n]$, where R is an integral domain. Let $S_1, \ldots, S_n$ be nonempty subsets of F and define $g_i(x_i) = \prod_{s \in S_i}(x_i - s)$. If $f(x_1, \ldots, x_n) = 0$ for all $(x_1, \ldots, x_n) \in S_1 \times \cdots \times S_n$, then there are polynomials $h_1, \ldots, h_n \in R[x_1, \ldots, x_n]$ such that $\deg(h_i) \leq \deg(f) - \deg(g_i)$ and*

$$f = \sum_{i=1}^{n} h_i g_i.$$

Proof. Write $g_i(x_i) = \prod_{s \in S_i}(x_i - s) = x_i^{t_i+1} - \sum_{j=0}^{t_i} g_{ij} x_i^j.$, and note that $x_i^{t_i+1} = \sum_{j=0}^{t_i} g_{ij} x_i^j$ for all $x \in S_i$.

Let $\overline{f}$ be the polynomial that results when each $x_i^{d_i}$ where $d_i > t_i$ is repeatedly replaced by $x_i^{d_i - t_i - 1} \sum_{j=0}^{t_i} g_{ij} x^j$ as much as possible. Note that $\overline{f}$ is obtained by subtracting products $h_i g_i$ from f, where the degree of h_i does not exceed $\deg(f) - \deg(g_i)$. Moreover, $\overline{f}(x_1, \ldots, x_n) = f(x_1, \ldots, x_n)$ for all $(x_1, \ldots, x_n) \in S_1 \times S_n$. Hence, $\overline{f}(x_1, \ldots, x_n) = 0$ for all such $(x_1, \ldots, x_n)$. Finally, by the brick lemma, $\overline{f} = 0$, completing the proof. □

I have called this theorem ideal decomposition because the equation for f reads $f \in (g_1, \ldots, g_n)$. In fact, in the multivariate case, the formulation is slightly stronger because it includes a nontrivial statement about the degrees involved. The statement about degrees is immediate in the single variable case, where it can clearly be improved to an equation. However, it is key to the power of the multivariate theory.

Theorem 5.2 (Inverse Bricks Theorem)**.** *Let $f \in R[x_1, \ldots, x_n]$, where R is an integral domain. Suppose the coefficient of $x_1^{t_1} \cdots x_n^{t_n}$ is nonzero, where $\sum_{i=1}^{n} t_i = \deg(f)$. Then if $S_1, \ldots, S_n$ are subsets of R with $|S_i| > t_i$, there exists a point $(x_1, \ldots, x_n) \in S_1 \times \cdots \times S_n$ such that*

$$f(x_1, \ldots, x_n) \neq 0.$$

Proof. Assume that $|S_i| = t_i + 1$ for all i. Suppose the result is false, and define $g_i(x_i) = \prod_{s \in S_i}(x_i - s)$. By the ideal decomposition theorem, there are polynomials $h_i \in F[x_1, \ldots, x_n]$ satisfying $\deg(h_j) \leq \sum_{i=1}^{n} t_i - \deg(g_j)$ such that

$$f = \sum_{i=1}^{n} h_i g_i.$$

The coefficient of $x_1^{t_1} \cdots x_n^{t_n}$ on the left is nonzero. However, the degree of $h_i g_i = h_i \prod_{s \in S_i}(x_i - s)$ is at most $\deg(f)$, and so any monomials of degree $\deg(f)$ in it are divisible by $x_i^{t_i+1}$. Because the monomial $x_1^{t_1} \cdots x_n^{t_n}$ violates this property for every i, its coefficient on the right is 0, the desired contradiction. □

Following a trend in the previous two results, the subsets of R^n involved in this generalization again take the form of a direct products. A distinction that

may be surprising is that direct products guaranteed to containing a non-zero are numerous. However, this multiplicity may be motivated by consideration of the explosion of potential maximal degree terms due to the definition of the degree of a multivariate polynomial.

It turns out that these multivariate generalizations are highly versatile in ways that are not anticipated by the single variable case. Developing a sense of this behavior is the goal of the exercises, to which we now turn.

5.1 Exercises

The theory described above, referred to as *combinatorial Nullstellensatz,*[1] provides a powerful tool for combinatorics and combinatorial number theory problems, as the following exercises will show. In fact, Alon proves these results and many more, so the interested reader is again referred his paper.

1. Prove the brick lemma.

2. (Chevalley-Warning theorem, Alon Theorem 3.1) Let p be a prime and let $P1, \ldots, P_m \in \mathbb{Z}/p\mathbb{Z}[x_1, \ldots, x_n]$ be m polynomials. If $n > \sum_{i=1}^{m} \deg(P_i)$ and the polynomials P_i have a common zero $(c_1, \ldots, c_n)$, then they have another common zero.

   a) Suppose the result is false and define

   $$f = \prod_{i=1}^{m} \left(1 - P_i(x_1, \ldots, x_n)^{p-1}\right) - \delta \prod_{j=1}^{n} \prod_{c \neq c_j} (x_j - c),$$

   where δ is such that $f(c_1, \ldots, c_n) = 0$. Show that $f(s_1, \ldots, s_n) = 0$ for all $(s_1, \ldots, s_n) \in (\mathbb{Z}/p\mathbb{Z})^n$.

   b) Find $s_1, \ldots, s_n \in \mathbb{Z}/p\mathbb{Z}$ such that $f(s_1, \ldots, s_n)$ is nonzero.

3. (Erdös-Ginzburg-Ziv theorem) For any positive integer n and any $2n-1$ integers, there are n whose sum is divisible by n.

   a) (Cauchy-Davenport, Alon Theorem 3.2) First establish the following lemma: if p is a prime and A, B are nonempty subsets of $\mathbb{Z}/p\mathbb{Z}$, then

   $$|A + B| \geq \min\{p, |A| + |B| - 1\}.$$

   Argue that it suffices to consider a counterexample $A + B \subset C$ where $|C| = |A| + |B| - 2$ and study the polynomial

   $$f = \prod_{c \in C} (x + y - c).$$

   b) Show that the case where n is prime follows from the Cauchy-Davenport lemma.

   c) Extend the result to all n by inducting on prime factors.

[1] The reader is invited to look up *Hilbert's Nullstellensatz* and establish the reason for the nomenclature.

4. (Alon Lemma 8.1) Let $A = (a_{ij})$ be an $n \times n$ matrix with coefficients in a field F. The *permanent* of A is defined as

$$Per(A) = \sum_{\sigma \in S_n} \prod_{i=1}^{n} a_{i,\sigma(i)},$$

where the sum runs over all permutations of n elements.

Suppose that $Per(A)$ is nonzero. Show that for any vector $b \in F^n$ and any family of sets $S_1, \ldots, S_n \subset F$, each of cardinality 2, there is a vector $x \in S_1 \times \cdots \times S_n$ such that Ax differs from b in every coordinate.

5. (Alon Theorem 4.1) Let p be a prime, $h \in \mathbb{Z}/p\mathbb{Z}[x_0, \ldots, x_k]$, and let $A_0, A_1, \ldots, A_k$ be nonempty subsets of $\mathbb{Z}/p\mathbb{Z}$. Define

$$\oplus_h \sum_{i=0}^{k} A_i = \{a_0 + \cdots + a_k : a_i \in A_i, h(a_0, \ldots, a_k) \neq 0\}.$$

Write $|A_i| = c_i + 1$ and define $m = -\deg(h) + \sum_{i=0}^{k} c_i$. Show that if the coefficient of $x_0^{c_0} \cdots x_k^{c_k}$ in

$$(x_0 + x_1 + \cdots + x_k)^m h(x_0, \ldots, x_k)$$

is nonzero as an element of $\mathbb{Z}/p\mathbb{Z}$, then

$$\left| \oplus_h \sum_{i=0}^{k} A_i \right| \geq m + 1.$$

6. (Alon Proposition 4.7) If p is prime and A and B are nonempty subsets of $\mathbb{Z}/p\mathbb{Z}$, then

$$|\{a + b : a \in A, b \in B, ab \neq 1\}| \geq \min\{p, |A| + |B| - 3\}.$$

7. (Alon Theorem 6.1) Let p be a prime and let $G = (V, E)$ be a loopless[2] graph with average degree bigger than $2p - 2$ and maximum degree at most $2p - 1$. Show that G contains a p-regular[3] subgraph.

8. (Alon Theorem 6.2) Let p be a prime and $G = (V, E)$ a graph with $|V| > d(p - 1)$ vertices where d is a positive integer. Show that there exists a nonempty subset $U \subset V$ of vertices such that the number of d-cliques that U intersects is 0 modulo p.

9. (IMO 2007/6) Let n be a positive integer. Consider

$$S = \{(x, y, z) | x, y, z \in \{0, 1, \ldots, n\}, x + y + z > 0\}$$

as a set of $(n + 1)^3 - 1$ points in three-dimensional space. Determine the smallest possible number of planes, the union of which contains S but does not include $(0, 0, 0)$.

[2] A *loop* is an edge from a vertex to itself.
[3] A *regular* graph is one in which all vertices have the same degree, p in this case.

References

[1] Bergeron, A. & Zhao, D., "Fermat a Frenicle," Sep 17, 2004, `http://www.cs.utexas.edu/users/wzhao/fermat2.pdf`, retrieved Dec 11 2007.

[2] "Pierre de Fermat," Wikipedia, Dec 11, 2007, `http://en.wikipedia.org/wiki/Pierre_de_Fermat`, retrieved Dec 11 2007.

[3] "Modular Elliptic Curves and Fermat's Last Theorem," *Annals of Mathematics* **141** (3), 443-551. 1995.

[4] Artin, M., "Algebra," Prentice Hall, 1991.

[5] Stanley, R. P., "Enumerative Combinatorics," Vol. 1, Wadsworth & Brooks/Cole, Monterey. 1986.

[6] "Hensel's lemma," Wikipedia, Dec 9, 2007, `http://en.wikipedia.org/wiki/Hensel%27s_lemma`, retrieved Dec 11 2007.

[7] "Isomorphism theorem," Wikipedia, Dec 7, 2007, `http://en.wikipedia.org/wiki/Isomorphism_theorem`, retrieved Dec 11 2007.

[8] Matthews, K. R., "A Generalization of Artin's Conjecture for Primitive Roots," *Acta Arith.* **29**, 113-146. 1976.

[9] Hooley, C., "On Artin's conjecture," *J. Reine Agnew. Math.* **225**, 209-220. 1967.

[10] Gupta, R. & Murty, M. R., "A remark on Artin's conjecture," *Invent. Math.* **78** (1), 127-130. 1984.

[11] Heath-Brown, D.R., "Artin's conjecture for primitive roots," *Quart. J. Math. Oxford Ser. (2)* **37**, 27-28. 1986.

[12] Alon, N., "Combinatorial Nullstellensatz," Combinatorics, Probability and Computing, **8** (1-2), 7-29. 1999.