1969

Reduced residue systems in the integers and other principal ideal domains

Paul E. Herrmann

The University of Montana

Let us know how access to this document benefits you.
Follow this and additional works at: https://scholarworks.umt.edu/etd

Recommended Citation
Herrmann, Paul E., "Reduced residue systems in the integers and other principal ideal domains" (1969). Graduate Student Theses, Dissertations, & Professional Papers. 8201.
https://scholarworks.umt.edu/etd/8201

This Thesis is brought to you for free and open access by the Graduate School at ScholarWorks at University of Montana. It has been accepted for inclusion in Graduate Student Theses, Dissertations, & Professional Papers by an authorized administrator of ScholarWorks at University of Montana. For more information, please contact scholarworks@mso.umt.edu.
REDUCED RESIDUE SYSTEMS IN THE INTEGERS
AND OTHER PRINCIPAL IDEAL DOMAINS

By

Paul E. Herrmann
B.S. University of Minnesota 1967

Presented in partial fulfillment of the requirements
for the degree of

Master of Arts for Teachers of Mathematics

UNIVERSITY OF MONTANA

1969

Approved by:

[Signatures]
Chairman, Board of Examiners
Dean, Graduate School

Date
Sept 25, 1969

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
ACKNOWLEDGEMENTS

I wish to thank Dr. Merle Manis for all the time and help he has given me in the preparation of this paper. I also wish to thank Dr. Mason Henderson for his suggestions and help in getting started, and Dr. William Ballard and Dr. Howard Reinhardt for their critical readings.

P. E. H.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Chapter I</td>
<td>A Classical Approach to Characteristics of Reduced Residue Systems.</td>
<td>3</td>
</tr>
<tr>
<td>Chapter II</td>
<td>A Generalization Through Algebra</td>
<td>11</td>
</tr>
<tr>
<td>Appendix</td>
<td></td>
<td>20</td>
</tr>
<tr>
<td>References</td>
<td></td>
<td>25</td>
</tr>
</tbody>
</table>

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
INTRODUCTION

This paper is divided into two parts. The first involves a discussion of some structural characteristics of reduced residue systems. The proofs are essentially classical and are centered around two theorems, parts of which are shown to be equivalent. Such well-known facts as the multiplicativity of the Euler $\phi$-function and the evaluation of $\phi(N)$ are seen to be provable as consequences of the two major theorems.

The second section of the paper is a reproving and generalizing of the ideas of the first section from an algebraic standpoint. Where the first theorems are proved for the integers, those of the second part are proved for principal ideal domains. Since the ring of integers is a principal ideal domain, the original theorems now are seen as special cases of their more general counterparts.

Although the paper was not intended as a comparison of the two approaches to the problems, the power of algebra cannot be overlooked as a tool for proving theorems in a less restricted setting. Secondly, one cannot fail to notice the interrelation of the two approaches and the value of being able to work from both.

One final comment should be made. Little of the material was obtained from books except for basic concepts; indeed, most of the material of Chapter II seems to lie in the realm of number theory folklore. Accordingly, the
main sources of information have been two series of conversations, one with Dr. Henderson and one with Dr. Manis.
CHAPTER I

A CLASSICAL APPROACH TO CHARACTERISTICS
OF REDUCED RESIDUE SYSTEMS

In this chapter theorems will be proved which deal with reduced residue systems. For convenience, a reduced residue system mod M will be abbreviated \( \mathbb{RR}_M \), and will be assumed to consist of the least positive representatives.

**Definition:** \( \mathbb{RR}_M = \{ x | 0 < x < M \text{ and } (x, M) = 1 \} \).

**Note 1:** In the sequel when a set of integers is referred to as \( \mathbb{RR}_M \), it will be assumed that the elements have been reduced mod M.

**Definition:** A complete residue system mod M = \( \{ x | 0 < x < M \} \).

**Definition:** A number A is called a preserving number relative to the modulus M, if, when A is added to each member of \( \mathbb{RR}_M \), \( \mathbb{RR}_M \) is again obtained.

Since \( a + M \equiv a \mod M \), it is clear that M itself is a preserving number for \( \mathbb{RR}_M \). On the other hand, \( M - 1 \) would not be a preserving number for \( \mathbb{RR}_M \), since \( 1 \in \mathbb{RR}_M \) while \( 1 + (M-1) \equiv 0 \mod M \) and \( 0 \notin \mathbb{RR}_M \).

**Theorem 1.1** Let \( M = \prod_{i \in I} p_i^{s_i} \), where the \( p_i \) are the distinct prime divisors of M. Then A is a preserving number for \( \mathbb{RR}_M \) if and only if A is a multiple of \( \prod_{i \in I} p_i \).

**Proof:** First assume that A is a multiple of \( \prod_{i \in I} p_i \), and \( r \in \mathbb{RR}_M ^- \). Clearly for all \( p_i \) such that \( i \in I \), \( p_i \mid A \) but \( p_i \nmid r \).
Therefore, \( p_1 \uparrow (A + r) \) and \((A + r, M) = 1\). Also if \( r_1, r_2 \in \mathbb{R}R_M \), then \( r_1 \neq r_2 \mod M \), which implies that \( A + r_1 \neq A + r_2 \mod M \). Therefore, by adding \( A \) to each element of \( \mathbb{R}R_M \), we obtain a reduced residue system \( \mod M \).

Now let \( A \) be a preserving number for \( \mathbb{R}R_M \) and let \( A = \prod_{j \in K} p_j^{t_j} \). It suffices to show that \( I = K \). Let 

\[ L = I \setminus (I \cap K) \]. If \( L \) is empty we are finished. Assuming \( L \neq \emptyset \), define \( c = \prod_{j \in L} p_j \). Let \( (cA, M) = v \). Then \( v \mid M \) so that \( v \) has factors \( p_1, i \in I \). However, \( I = K \cup L \) and thus for all \( p_1, i \in I \), we have \( p_1 \mid c \) or \( p_1 \mid A \). But \( p_1 \mid c \Rightarrow i \in L \Rightarrow i \notin K \Rightarrow p_1 \uparrow A \). Also, \( p_1 \mid A \Rightarrow i \in K \Rightarrow i \notin L \Rightarrow p_1 \uparrow c \). Therefore, \( i \in I \Rightarrow p_1 \uparrow (c - A) \Rightarrow v = 1 \). Thus there exists \( q \in \mathbb{R}R_M \) such that \( q \equiv c - A \mod M \). But then \( A + q \equiv A + (c - A) \equiv c \mod M \), and \( A \) is not a preserving number. It must therefore be false that \( L \neq \emptyset \); i.e., \( I = K \).

**Theorem 1.2** Let \( \mathbb{R}R_M \) and \( \mathbb{R}R_N \) be reduced residue systems such that \( M \mid N \). Suppose \( x \in \mathbb{R}R_N \). Then there exists \( z \in \mathbb{R}R_N \) such that \( x \equiv z \mod M \).

**Proof:** Let \( N = \prod_{i \in I} p_i^{s_i} \) and \( K = \{ i \mid i \in I \text{ and } p_i \mid x \} \). Now if \( (x, N) = 1 \), we are finished. Therefore, assume \( (x, N) \neq 1 \), that is \( K \neq \emptyset \). Define \( c \) by \( c = \prod_{i \in K} p_i^{s_i} \). Consider \( x + N/x \).

By construction we have \( (c, M) = 1 \) and \( (x, N/c) = 1 \). Also we have for any \( p_1 \) such that \( i \in I \), either \( p_1 \mid x \) and \( p_1 \uparrow N/c \) or \( p_1 \mid N/c \) and \( p_1 \uparrow x \). Therefore \( (x + N/c, N) = 1 \). But \( M \mid N \).
and \((M, c) = 1\) implies \(M \mid N/c\). Letting \(z = x + N/c\), we have \(z \equiv x \mod M\), finishing the proof.

For an example of the above, let \(M = 12\) and \(N = 2100\). Now \(12 \mid 2100\) and \(5 \in \text{RR}_{12}\) while \(5 \notin \text{RR}_{2100}\). To find \(z\) we proceed as the theorem indicates. We see that \(K = \{3\}\) where \(p_3 = 5\), \(c = 5^2\), \(x + N/c = 5 + \frac{2100}{25} = 5 + 84 = 89\). Thus 89 is the required \(z\) since \(89 \in \text{RR}_{2100}\) and \(5 \equiv 89 \mod 12\).

This theorem shows us that a reduced residue system mod \(N\) "contains" a reduced residue system mod \(M\). In the next theorem, we will get a more exact idea of the structure of \(\text{RR}_N\).

**Theorem 1.3** If \(N = \prod_{i=1}^{n} p_i^{s_i}\), then for any \(j\), such that \(1 < j \leq n\), a reduced residue system mod \(N\) contains exactly

\[\prod_{i=1}^{n} \frac{(p_i - 1)(p_i - 1)}{(p_j - 1)}\]

elements which are congruent to \(x \mod p_j\), \(1 \leq x \leq p_j - 1\).

**Proof:** First we prove that if \(J = \{x, x + p_j, x + 2p_j, \ldots, x + (N - p_j)\}\), then for any \(p_i\), \(i \neq j\), each \(p_i\) consecutive elements of \(J\) form a complete residue system mod \(p_i\). To see this let \(x + np_j\) and \(x + mp_j\) be two elements of \(J\) occurring in \(p_i\) consecutive elements. We need only show that \(x + np_j \neq x + mp_j \mod p_i\). Suppose then that \(x + np_j \equiv x + mp_j \mod p_i\). Therefore \(p_i \mid (m - n)p_j\) and \(p_i \mid p_j\) or \(p_i \mid (m - n)\). But \(p_i\) and \(p_j\) are distinct primes so \(p_i \nmid p_j\). Also \(|m - n| < p_i\) which
implies that \( p_1 \nmid (m - n) \). Therefore the proposition holds. Note that \( J \) contains all those elements we wish to count plus many which are not relatively prime to \( N \).

We now choose \( x \) such that \( 1 \leq x \leq p_j - 1 \) and proceed to count the elements of \( J \) which are relatively prime to \( N \). From the above we know that of each \( p_i \) consecutive elements in \( J \), one and only one is congruent to zero mod \( p_i \) and therefore not relatively prime to \( N \). These elements must be deleted from \( J \) for all \( p_i \). Note that \( J \) has \( N/p_i \) elements so that there are exactly \( N/p_i p_j \) complete residue systems mod \( p_i \) and thus exactly \( N/p_i p_j \) elements in \( J \) which must be subtracted from \( N/p_j \) to obtain the number of elements in \( J \) which are not divisible by \( p_i \). If this subtraction is carried out for all \( p_i, i \neq j \), we will be left with the number of elements in \( J \) which are relatively prime to all \( p_i \), and thus to \( N \). We count, being careful not to subtract the same element more than once. The number we obtain is:

\[
\frac{N}{p_j} - \frac{1}{p_j} \sum_{i \neq j} \frac{N}{p_i} + \frac{1}{p_j} \sum_{i \neq k} \frac{N}{p_i p_k} - \cdots + \frac{1}{p_j} \frac{N}{\prod_{j=1}^{k} \frac{p_i}{p_j}} - \sum_{i \neq j} \frac{1}{p_j} \left( \frac{\prod_{i=1}^{n} (p_i^{-1})}{(p_j - 1)} \right) (p_j - 1) \left\{ \frac{\prod_{i=1}^{n} p_i}{p_j} - \frac{1}{p_j} \sum_{i \neq j} \frac{\prod_{i=1}^{n} p_i}{p_i} + \cdots + 1 \right\}
\]

\[
= \frac{\prod_{i=1}^{n} (p_i^{-1})}{(p_j - 1)} (p_j - 1) \prod_{i \neq j}^{n} (p_i - 1) = \frac{\prod_{i=1}^{n} (p_i^{-1}) (p_i - 1)}{(p_j - 1)}
\]

From Theorem 1.3 we know exactly how many elements there are in \( \mathbb{R}^N \) which are congruent to \( x \) mod \( p_j \). Since
this number is the same for each \( x \) such that \( 1 \leq x \leq p_j - 1 \) we know that \( RR_N \) contains at least \( \prod_{i=1}^{n} (p_i^{s_i-1})(p_i - 1) \) elements. Also, if \( x = 0 \mod p_j \), then \( x \not\in RR_N \) so that \( RR_N \) contains at most \( \prod_{i=1}^{n} (p_i^{s_i-1})(p_i - 1) \) elements. We have thus proved the following corollary.

**Corollary 1.4** If \( N = \prod_{i=1}^{n} p_i^{s_i} \), then \( RR_N \) contains exactly \( \prod_{i=1}^{n} (p_i^{s_i-1})(p_i - 1) \) elements. In standard notation, \( \phi(N) = \prod_{i=1}^{n} (p_i^{s_i-1})(p_i - 1) \).

Having arrived at this corollary without use of the fact that \( \phi \) is multiplicative, it is easy to see that Theorem 1.3 and Corollary 1.4 establish the following:

**Corollary 1.5** \( \phi \) is a multiplicative function.

**Proof:** Let \( M = \prod_{i \in I} p_i^{s_i} \) and \( N = \prod_{i \in K} p_i^{t_i} \). Then if \( (M, N) = 1 \) we have \( I \cap K \) empty so that \( MN = \prod_{i \in I} p_i^{s_i} \prod_{i \in K} p_i^{t_i} \) and \( \phi(MN) = \prod_{i \in I} (p_i^{s_i-1})(p_i - 1) \prod_{i \in K} (p_i^{t_i-1})(p_i - 1) = \phi(M)\phi(N) \).

The following theorem is a generalization of Theorem 1.3; part of the detail is left out since the counting idea is similar to that done in Theorem 1.3.

**Theorem 1.6** Let \( N \) be any divisor of \( M \) such that \( N > 1 \). Then \( RR_N \) can be partitioned into sets, each of which is a reduced residue system mod \( M \).

**Proof:** Let \( M = \prod_{i=1}^{n} p_i^{s_i} \), \( N = \prod_{i=1}^{n} p_i^{t_i} \) where \( s_i \geq t_i \geq 0 \). Now define \( u_i = \max(t_i, 1), K = \{i | t_i = 0\} \). We now write
M = R·Q where $R = \prod_{i=1}^{n} p_i^{u_i}$, $Q = \prod_{i=1}^{n} p_i^{s_i-u_i}$. Now the proof can be divided into two cases.

**Case i:** If $K$ is empty, then $R = N$, and we have $(x, N) = 1$ iff $(x, M) = 1$. Therefore, $RR_N$ constitute the first $\phi(N)$ elements occurring in $RR_M$. According to Theorem 1.1, $N$ is a preserving number for both $RR_N$ and $RR_M$. Then by adding $N$ to $RR_N$ and not reducing mod $N$, we obtain the next $\phi(N)$ elements of $RR_M$, namely those between $N$ and $2N$. Continuing in this manner, we find all elements of $RR_M$ by looking at copies of $RR_N$ (non-reduced mod $N$). By the definition of $Q$ we see that there are precisely $Q$ such copies. It might be natural to ask if this agrees with Corollary 1.4 and we see that it actually does; i.e., $\phi(N)Q = \phi(M)$.

**Case ii:** $K$ is not empty. Then $R/N = \prod_{i \in K} p_i$. Let $x \in RR_N$ and consider a set $J$ as in Theorem 1.3; $J = \{ x, x + N, x + 2N, \ldots x + (R - N) \}$. Here, as in Theorem 1.3 we can count the elements in $J$ which are relatively prime to $N$, obtaining:

$$
\frac{R/N - l/N}{\sum_{i \in K} R/p_i + l/N} \sum_{i,j \in K} R/p_i p_j - \ldots \pm l/N \cdot R/ \prod_{i \in K} p_i
$$

$$
= \prod_{i \in K} p_i - \sum_{i \in K} \prod_{i \neq j} p_i/p_j + \sum_{i,j \in K} \prod_{i \neq j} p_i/p_i p_j - \ldots \pm 1
$$

$$
= \prod_{i \in K} (p_i - 1).
$$

Thus each element of $RR_N$ occurs $\prod_{i \in K} (p_i - 1)$ times in $RR_R$. 

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
But now we have a situation exactly like case i. That is, \( RR_N \) is made up of \( Q \) copies of \( RR_R \). Each \( RR_R \) has \( \prod_{i \in K} (p_i - 1) \) copies of \( RR_N \). Thus \( RR_N \) consists of exactly \( Q \prod_{i \in K} (p_i - 1) \) sets, each of which is a reduced residue system mod \( N \).

Again it is natural to compute \( \phi(M) \) to see if there is agreement between the above and Corollary 1.4. Computing:

\[
\phi(N) \cdot Q \cdot \prod_{i \in K} (p_i - 1) = \prod_{i \in K} (p_i^{t_i - 1})(p_i - 1) \cdot Q \cdot \prod_{i \in K} (p_i - 1)
\]

\[
= \prod_{i \in K} (p_i^{t_i - 1}) \cdot \prod_{i \in K} (p_i^{s_i - 1}) \cdot \prod_{i=1}^n (p_i - 1)
\]

\[
= \phi(M). \quad \text{Thus we see that this counting of } \phi(M) \text{ does agree with Corollary 1.4.}
\]

**Theorem 1.7** The necessary condition of Theorem 1.1 is equivalent to a weaker statement of Theorem 1.2. Precisely:

(a) \( A \) is a preserving number for \( RR_M \) only if \( A \) is a multiple of all the primes occurring in \( M \) is equivalent to (b) If \( p \) is any prime which divides \( M \), then for all \( x, 1 \leq x \leq p - 1 \), there exists \( z \) such that \( z \in RR_M \) and \( x \equiv z \mod p \).

**Proof, part i:** First we will show that (b) implies (a).

Let \( N \) be any number such that there is some prime \( p \) such that \( p | M \) but \( p \nmid N \). Since \( p \nmid N \), \( N \equiv x \mod p \) for some \( x \) such that \( 1 \leq x \leq p - 1 \). Now by our assumption, (b), there exists some \( z \in RR_M \) such that \( z \equiv p - x \mod p \). Therefore \( N + z \equiv x + (p - x) \equiv 0 \mod p \), i.e., \( N + z \not\in RR_M \). Thus
N is not a preserving number for $\mathcal{RR}_M$. Thus $(b) \Rightarrow (a)$.

**part ii:** $(a) \Rightarrow (b)$. Let $M = \prod_{i=1}^{n} p_i^{s_i}$ and let $p_j$ be any prime divisor of $M$. Define $P_j = \prod_{i=1}^{n} p_i/p_j$ and let $S = \{P_jn|n = 1, 2, \ldots, p_j - 1\}$. Since $p_j \nmid P_j$, $S$ is a reduced residue system mod $p_j$. Now let $x$ be any number such that $1 \leq x \leq p_j - 1$. Since $S$ is a reduced residue system mod $p_j$, there exists $k$ such that $1 \leq k \leq p_j - 1$, and $P_jk \equiv p_j - x \mod p_j$. Also since $p_j \nmid P_j$, we know that all elements of $S$ fail to be preserving numbers. Therefore, there exists $z \in \mathcal{RR}_M$ such that $z + P_jk \notin \mathcal{RR}_M$. But for all $i \neq j$, we have $p_i \nmid P_jk$ and $p_i \nmid z$, implying that $p_i \nmid (z + P_jk)$. Therefore, $p_j \nmid z + P_jk$ which implies that $z + P_jk \equiv 0 \mod p_j$. But $p_j - x \equiv P_jk \Rightarrow z + (p_j - x) \equiv 0 \mod p_j \Rightarrow z \equiv x \mod p_j$. We have thus found $z$ such that $z \in \mathcal{RR}_M$ and $z \equiv x \mod p_j$. From (i) and (ii) we have the desired equivalence.

In order to look at some of these properties in a more general setting, we now switch to an algebraic approach. Some of the theorems which are basic are proved in the Appendix and should be reviewed prior to Chapter II.
CHAPTER II
A GENERALIZATION THROUGH ALGEBRA

In this chapter we will be working with rings, which we will assume to be commutative.

Definition: A ring R is called a principal ideal domain if it satisfies the following:

1) R has a multiplicative identity 1, 1 ≠ 0
2) If A is an ideal of R, A = {at | t ∈ R} for some a in R
3) R has no zero divisors

Examples of principal ideal domains include the integers, all fields, and polynomials over a field. In the following R is a principal ideal domain and U(R) is the set of units of R.

Definition: a ∈ R, then (a) = {at | t ∈ R}

Note 1: (a) = (b) iff a = ub for some unit u of R. Certainly if a = 0, then b = 0 and a = 1 · b. Suppose a ≠ 0, b ≠ 0. Now b ∈ (b) ⇒ b ∈ (a) ⇒ b = at; similarly a = bs.
Thus a = at ⇒ a(1 - st) = 0 ⇒ 1 = st. That is, s and t are units. In particular if s is any unit, then (1) = (s) = R.

Definition: (a, b) = d where (d) = {ta + sb | t, s ∈ R}.

We see that this definition has meaning since (a) + (b) = {ta + sb | t, s ∈ R} is an ideal, hence (d) for some d ∈ R.

The definition, however, is only given "up to a unit" since (d) may equal (d'). In this latter case, d = ud'
for some unit u. Note \( d = ax + by \) for some \( x, y \in R \).

**Note 2:** If \( (a, b) = d \), then \( (a) = (d) \) since \( at = at + Ob \in (d) \). Similarly \( (b) \subseteq (d) \).

**Definition:** \( a \in R \) is called prime if \( (a) \) is a prime ideal of \( R \); i.e. \( xy \in (a) \Rightarrow x \in (a) \) or \( y \in (a) \). \( a \) is a proper prime if \( a \) is prime and \( (a) \neq (0) \) and \( (a) \neq (1) \).

**Definition:** \( a \) divides \( b \), written \( a \mid b \) iff \( (b) \subseteq (a) \).

**Definition:** \( \sqrt{A} \), the radical of \( A \), for an ideal \( A \), is equal to \( \cap \{ (p) \mid p \text{ is a prime and } A \subseteq (p) \} \).

**Definition:** Let \( A \) and \( B \) be ideals. Then \( B : A = \{ x \mid xy \in B \text{ for all } y \in A \} \). ("B:A" is read B divided by A).

**Lemma 2.1** If \( \sqrt{(b)} : (a) = (q) \) and \( b \neq 0 \), then \( (a, q) = 1 \). Note that \( q \neq 0 \) since \( b \in \sqrt{(b)} \).

**Proof:** Let \( (a, q) = d \); then \( a = dh \) and \( q = dt \). Now \( ta = t(dh) = (td)h = (dt)h = qh \) so that \( ta^2 = (qa)h = q(ah) \in \sqrt{(b)} \). \( ta^2 \in \sqrt{(b)} \Rightarrow ta^2 \in (p) \) for all \( p \) such that \( (b) \subseteq (p) \). This implies \( t \in (p) \) or \( a \in (p) \) for all prime \( p \); therefore \( ta \in (p) \) for all \( p \) and thus \( ta \in \sqrt{(b)} \). Now \( ta \in \sqrt{(b)} \Rightarrow t(a) \subseteq \sqrt{(b)} \Rightarrow t \in (q) \), so \( sq = t \) for some \( s \) and thus \( q = (ds)q \) which proves \( d \) is a unit and therefore \( (d) = (1) \).

**Lemma 2.2** If \( b \neq 0 \), \( (a, x) = 1 \), and \( (q) = \sqrt{(b)} : (a) \), then \( (a + qx, b) = 1 \).

**Proof:** Let \( (a + qx, b) = e \). If \( (e) \neq (1) \), then there is a proper prime \( p \) with \( (e) \subseteq (p) \). See Appendix, Corollary G.1. Therefore \( (a + qx, b) \in (p) \).
Case i: $q \in (p) \Rightarrow qx \in (p) \Rightarrow (a + qx) - qx = a \in (p)$; but $(a, q) = 1$ which implies $(a) + (q) = (1) \subseteq (p)$. This is a contradiction since $p$ is a proper prime.

Case ii: $q \not\in (p)$; since $q \cdot a \in \sqrt{(b)} \subseteq (p)$, we have $a \in (p)$. But then $qx = (a + qx) - a \in (p)$; now $qx \in (p)$ and $q \not\in (p) \Rightarrow x \in (p) \Rightarrow (a) + (x) = (1) \subseteq (p)$. Again we have a contradiction. Therefore $(a) = (1)$.

Lemma 2.3 Let $f_R \rightarrow R/(a)$ be the natural map. Then $f(b)$ is a unit of $R/(a)$ iff $(a, b) = 1$.

Proof: $(a, b) = 1 \Rightarrow ax + by = 1$ for some $x, y \in R$. Therefore $f(ax + by) = f(by) = f(b)f(y)$. Also $f(ax + by) = f(1) \Rightarrow f(b)f(y) = f(1)$ so that $f(b)$ is a unit. Now suppose $f(b)$ is a unit; i.e. $f(b)f(x) = f(1)$ for some $x$. But then $f(bx) - f(1) = f(bx - 1) = f(0)$; thus $bx - 1 = -ya$ for some $y \in R$ or $bx + ya = 1$. Now $bx + ya = 1 \Rightarrow 1 \in (a) + (b) \Rightarrow (a) + (b) = (1)$; i.e. $(a, b) = 1$.

Theorem 2.4 If $x|b$, then the natural map $f$ takes $U(R/(b))$ onto $U(R/(x))$.

Proof: To see that units map into units, we note that from Lemma 2.3 that $a + (b)$ is a unit of $R/(b)$ iff $(a, b) = 1$. Also we know that $x|b$; i.e. $(b) \subseteq (x)$. Now $(a, b) = 1 \Rightarrow (a) + (b) = (1) \Rightarrow (a) + (x) = (1)$, or $(a, x) = 1$. But $(a, x) = 1$ iff $a + (x)$ is a unit of $R/(x)$.

Now let $(a, x) = 1$; i.e. $a + (x)$ is a unit of $R/(x)$. Let $(q) = \sqrt{(b)}$: $(a)$ and let $c = a + qx$. Then from Lemma 2.2, $(c, b) = 1$ and $f(c) = a + qx + (x) = a + (x)$. Thus
we have found \( c \), a unit of \( R/(b) \) such that \( f(c) = a + (x) \); i.e. \( f \) maps \( U(R/(b)) \) onto \( U(R/(b)) \). Letting \( R = \mathbb{Z} \) (the integers) we see that this theorem is the analog of Theorem 1.2.

Since the units of both \( R/(b) \) and \( R/(x) \) form multiplicative groups, \( f \) can be viewed as a group homomorphism. Thus for each unit \( a + (x) \) of \( R/(x) \), the set of inverse images is \( sK \) where \( s \) is one inverse image and \( K \) is the kernel of \( f \). We have proved the following corollary.

**Corollary 2.5** If \( x|b \), then \( U(R/(b)) \) can be partitioned into sets \( t_i \), such that for each \( u_i \) of \( R/(x) \) there is a set \( t_i \) such that \( f: t_i \rightarrow u_i \), and each \( t_i \) contains the same number of elements, namely the number of elements in the kernel of \( f \).

The preceding corollary is the analog of Theorem 1.6.

**Theorem 2.6** Consider \( V = R/(n) \). Let \( \overline{x} = x + (n) \).

Then \( \{ \overline{x} | \overline{x} + U(V) = U(V) \} = \{ \overline{x} | p| x \ \text{for all} \ p \ \text{such that} \ p|n \} \).

**Proof:** Suppose there exists \( p \) such that \( p|n \) but \( p\nmid x \). Then \( f(\overline{x}) \) is a unit of \( R/(p) \) since \( f(\overline{x}) = x + (n) \) + \( (p) = x + (p) = f(x) \), which is a unit by Lemma 2.3. Then by Theorem 2.4 there exists \( \overline{y} \in U(V) \) such that \( f(\overline{y}) = f(\overline{x}) \). But then \( f(\overline{x} - \overline{y}) = \overline{0} \) is not a unit of \( R/(p) \). This implies \( \overline{x} - \overline{y} \) is not a unit of \( R/(n) \) and therefore \( \overline{x} \) is not in \( \{ \overline{x} | \overline{x} + U(V) = U(V) \} = S_1 \). Letting \( S_2 = \{ \overline{x} | p| x \ \text{for all} \ p \ \text{such that} \ p|n \} \), we have \( S_1 \subseteq S_2 \).
Now suppose \( x \in S_2 \), and let \( \bar{u}_i = u_i + (n) \) be a unit of \( R/(n) \). Since \( \bar{u}_i \) is a unit of \( R/(n) \), we have \((u_i, n) = 1\). But now \((u_i + x, n) = 1\) since if \( \bar{x} = x + (n) \in S_2 \), then all \( p \) which divide \( n \) also divide \( x \). But now \( u_i + x + (n) = u_i + (n) + x + (n) = \bar{u}_i + \bar{x} \) is a unit of \( R/(n) \). Therefore \( S_2 \subseteq S_1 \). By looking at \( S_1 \) as the set of preserving numbers, we see this theorem as the analog of Theorem 1.1.

**Theorem 2.7** \( \sqrt{(n)} = \{x | x + U(R/(n)) \in U(R/(n))\} \).

**Proof:** Let \( x \) be an element of the second set. Then by Theorem 2.6 we have \( p | n \Rightarrow p | x \) and thus \( (n) = (p) \Rightarrow (x) \in (p) \). This implies \( (x) \subseteq \bigcap \{p | (n) = (p), p \text{ prime}\} \). Therefore \( (x) \subseteq \sqrt{(n)} \) or \( x \in \sqrt{(n)} \). Now let \( y \in \sqrt{(n)} \), and let \( \bar{u} \) be a unit of \( R/(n) \). Then by Corollary G.4 of the Appendix we have that \( \bar{y} + \bar{u} \) is again in \( U(R/(n)) \). Therefore \( y \in \{x | x + U(R/(n)) \subseteq U(R/(n))\} \).

If \( R_1, R_2, \ldots, R_n \) are rings, then \( R_1 \oplus R_2 \oplus \ldots \oplus R_n \) = \( \{(x_1, x_2, \ldots, x_n) | x_i \in R_i\} \) is also a ring where \( (x_1, x_2, \ldots, x_n) + (y_1, y_2, \ldots, y_n) = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n) \) and \( (x_1, x_2, \ldots, x_n) \cdot (y_1, y_2, \ldots, y_n) = (x_1 y_1, x_2 y_2, \ldots, x_n y_n) \). This ring is denoted by \( \bigoplus_{i=1}^{n} R_i \) and is read "direct sum of the \( R_i \)". We are interested in the following case. Let \( (a) \) and \( (b) \) be ideals of \( R \) such that \( (a) \cap (b) = (ab) \), i.e. \((a, b) = 1\). Then \( R/(a) \) and \( R/(b) \) are rings so that \( R/(a) \oplus R/(b) \) is also a ring.
Theorem 2.8  Let (a) and (b) be ideals of R such that (a, b) = 1. Then R/(ab) \cong R/(a) \oplus R/(b).

Proof: If R is a ring, we have ring homomorphisms \( f_1 \) and \( f_2 \) from R onto R/(a) and R/(b) respectively. Then we can define f from R into R/(a) \oplus R/(b) by \( f(x) = (f_1(x), f_2(x)) \). Clearly f is a homomorphism since

\[
\begin{align*}
f_2(x + y) &= (f_1(x) + f_1(y), f_2(x) + f_2(y)) = (f_1(x), f_2(x)) \\
&\quad + (f_1(y), f_2(y)) = f(x) + f(y).
\end{align*}
\]

Similarly \( f(xy) = f(x)f(y) \).

Thus we know from Theorem D of the appendix that R/Kernel f is isomorphic to the image of f. We must show f is onto. Let \((x_1 + (a), x_2 + (b)) \in R/(a) \oplus R/(b)\). We need n such that \( f(n) = (x_1 + (a), x_2 + (b)) \); i.e. \( n \equiv x_1 \mod a \) and \( n \equiv x_2 \mod b \). Since \( (a, b) = 1 \), we have \( s, t \in R \) such that \( as + bt = 1 \); equivalently \( as + bt \equiv 1 \mod a \) and \( as + bt \equiv 1 \mod b \) \( \Rightarrow \) \( as_1 + btx_1 = x_1 \mod a \) and \( as_2 + btx_1 = x_2 \mod b \).

We have found n; therefore f is onto.

Now the kernel of \( f = \{x | f_1(x) = 0 \text{ and } f_2(x) = 0\} \).

Thus kernel f = kernel \( f_1 \cap \) kernel \( f_2 = (a) \cap (b) = (ab) \) since \( (a, b) = 1 \). Therefore R/(ab) \cong R/(a) \oplus R/(b).

Corollary 2.9  Let \( p_1, p_2, \ldots, p_n \) be n distinct primes of R. Then \( R/(\prod_{i=1}^{n} p_i^{s_i}) \cong \bigoplus_{i=1}^{n} R/(p_i^{s_i}) \).

Proof: The proof is made by induction; we note that the ideals \((p_i^{s_i})\) satisfy the conditions of Theorem 2.8 for any pair. Therefore \( R/(p_1^{s_1}p_2^{s_2}) \cong R/(p_1^{s_1}) \oplus R/(p_2^{s_2}) \). Now suppose \( R/(\prod_{i=1}^{k} p_i^{s_i}) \cong \bigoplus_{i=1}^{k} R/p_i^{s_i} \). Then since \((\prod_{i=1}^{k} p_i^{s_i}) = 1\),
we again have \[ \frac{R}{(\prod_{i=1}^{k} p_i^{s_i})} \cong \frac{R}{(\prod_{i=1}^{k} p_i^{s_i}) \oplus \frac{R}{(p_{k+1}^{s_{k+1}})}} \]
\[ \cong \bigoplus_{i=1}^{k} \frac{R}{(p_i^{s_i})} \oplus \frac{R}{(p_{k+1}^{s_{k+1}})}. \]
Therefore \[ \frac{R}{(\prod_{i=1}^{k+1} p_i^{s_i})} \cong \bigoplus_{i=1}^{k+1} \frac{R}{(p_i^{s_i})}, \]
completing the proof.

If \( R = \mathbb{Z} \) in the previous theorem we have a proof of
The Chinese Remainder Theorem. It is also true that Theorem
2.4 and Theorem 2.7 could be obtained by a close inspection
and analysis of the decomposition of Corollary 2.9. The
thing to notice is that \( R/(n) \) is a finite direct sum of rings,
each having a unique prime ideal.

Corollary 2.10 \( \phi(N) = \prod_{i=1}^{n} (p_i^{s_i-1})(p_i - 1) \).

Proof: Letting \( R = \mathbb{Z} \) in the above we have immediately that
\[ \frac{\mathbb{Z}}{(\prod_{i=1}^{n} p_i^{s_i})} \cong \bigoplus_{i=1}^{n} \frac{\mathbb{Z}}{(p_i^{s_i})}. \]
Therefore the number of
units in \( \frac{\mathbb{Z}}{(\prod_{i=1}^{n} p_i^{s_i})} \) is the same as the number of units
in \( \bigoplus_{i=1}^{n} \frac{\mathbb{Z}}{(p_i^{s_i})} \), and units correspond directly to elements
of \( RR_{N} \). For \((x_1, x_2, \ldots, x_n)\) to be a unit of \( \bigoplus_{i=1}^{n} \frac{\mathbb{Z}}{(p_i^{s_i})} \),
we must have a unit of \( \frac{\mathbb{Z}}{(p_i^{s_i})} \) at the \( i \)th component for all
\( i \). Thus the number of units can be counted by counting the
number of possible units at each component. But since \( p_i \)
is a prime for all \( i \), the number of units of \( \frac{\mathbb{Z}}{(p_i^{s_i})} \) is
\[ p_i^{s_i} - p_i^{s_i}/p_i = p_i^{s_i-1}(p_i - 1). \] Multiplying we find the total
number of units to be \( \prod_{i=1}^{n} (p_i^{s_i-1})(p_i - 1) \).

Note that although the previous corollary was proved
for some \( N = \prod_{i=1}^{n} p_i^{s_i} \), an integer, it is in fact true that in any principal ideal domain an element \( a \) can be written as \( \prod_{i=1}^{n} p_i^{s_i} \) for distinct primes \( p_i \). This counting of \( \phi(N) \), together with Corollary 2.5 provides us with a counting of the kernel of \( f \) in Theorem 2.4 and thus with the analog of Theorem 1.3; i.e. \( (\phi(N)/\text{order of the kernel}) = \) the number of elements in any \( t_1 \) of Corollary 2.5.

**Example:** Let \( F \) be a finite field with \( |F| = q^n = s \), where \( q \) is some prime. Let \( R = F[x] \) and let \( p = r(x) \) be a prime of \( R \); i.e. \( r(x) \) is irreducible. Let \( f:R \longrightarrow R/(p) \) be the natural map. Then \( f(t) \in U(R/(p)) \) iff \( (t, p) = 1 \).

If degree \( t \geq \) degree \( p \), and \( (t, p) = 1 \), we have \( t = hp + r \) where degree \( r < \) degree \( p \). Also \( (r, p) = 1 \) and \( f(t) = f(r) \).

Therefore the units of \( R/(p) \) are images of polynomials of degree less than the degree of \( p \). Thus \( \phi(p) \) becomes the number of polynomials of degree less than degree \( p \) and relatively prime to \( p \). Of course since \( p \) is prime we have \( \phi(p) = s^d - 1 \) where \( d = \) degree of \( p \). (The number \( s^d - 1 \) is obtained by considering the \( s \) possible coefficients at \( d \) places and then subtracting \( 1 \) for the polynomial 0).

Continuing we can look at \( \phi(p^e) \) which is the number of polynomials of degree less than \( p^e \) and relatively prime to \( p^e \). This can be obtained by considering all polynomials of degree less than \( de \) and subtracting those of degree less than \( de \) which are divisible by \( p \). We obtain \( s^{de} - s^{de}/s^d \) = \( s^{de-d}(s^d - 1) \). We see this formula is an analog of
Corollary 2.10 where $s^d = p_1$ and $e = s_1$. 

APPENDIX

The following ideas are considered essential for an understanding of Chapter II.

**Theorem A.** \( a \in \mathbb{R} \iff a + (b) \in U(\mathbb{R}/(b)). \)

**Proof:** Let \( a \in \mathbb{R} \implies (a, b) = 1 \implies \) there exist \( x, y \in \mathbb{R} \) such that \( ax + by = 1 \) which implies \( ax \equiv 1 \mod b \) and thus \( (a + (b))(x + (b)) \equiv 1 \mod b; \) i.e., \( a + (b) \) is a unit. Now suppose \( (a + (b))(x + (b)) \equiv 1 \mod b. \) We then have \( ax + (b) \equiv 1 \mod b \implies ax - 1 \equiv 0 \mod b \implies ax - 1 = -yb \) for some \( y \) \( \implies ax + yb = 1; \) i.e., \( (a, b) = 1 \) or \( a \in \mathbb{R}. \)

**Theorem B.** \( \mathbb{R}_b, \) or equivalently the units of \( \mathbb{R}/(b), \) form a multiplicative group.

**Proof:** By their very definition units have multiplicative inverses. Also \( 1 \cdot 1 = 1 \) so that \( 1 \) is certainly a unit.

Multiplication of units is associative since multiplication is associative in the original ring. Finally, let \( a \) and \( b \) be units; then \( ab \cdot b^{-1} a^{-1} = 1 \) so that \( ab \) is again a unit.

**Theorem C.** Let \( G \) and \( H \) be groups and let \( f : G \to H \) be a group homomorphism. Then the set \( \{x | f(x) = h \text{ for } h \in H \} = \{kx_1 | k \in \text{kernel } f \text{ and } x_1 \text{ is some element of } G \text{ such that } f(x_1) = h\}. \)

**Proof:** Suppose \( f(x_1) = h; \) then \( f(x_1 k) = f(x_1)f(k) = h \cdot 1 = h. \)

Now suppose \( f(z) = h = f(x_1). \) Then \( f(z) \cdot f(x_1)^{-1} = 1 \implies f(zx_1^{-1}) = zx_1^{-1} \) is in the kernel of \( f \implies zx_1^{-1} = k \implies z = kx_1. \)

**Theorem D.** Let \( R \) and \( S \) be rings and let \( f \) be a ring
homomorphism from $R$ onto $S$. Let $K$ be the kernel of $f$. Then $R/K \cong S$.

**Proof:** Since $K$ is an ideal, there is a natural map $h$ from $R$ onto $R/K$, $h(r) = K + r$. We define the map $g$ by $g(K + r) = f(r)$. We shall show $g$ is an isomorphism from $R/K$ onto $S$. First we must be sure that $g$ is well defined. Suppose $K + r = K + s$; then $s \in K + r \Rightarrow s = k + r$ for some $k \in K$. But then by Theorem $C$ we have that $g(K + r) = f(r) = f(k + r)$ $= f(s) = g(K + s)$ so that our function is well defined. Now let $v \in S$. Since $f$ is onto, there exists $r \in R$ such that $f(r) = v$. But then $g(K + r) = v$ so that $g$ is onto. To see that $g$ is a homomorphism, consider $g[(K + r) + (K + r')]$ $= g[K + (r + r')] = f(r + r') = f(r) + f(r') = g(K + r) + g(K + r')$. Finally we must show that $g$ is one-to-one. Suppose $g(K + a) = 0 \Rightarrow f(a) = 0 \Rightarrow a \in K \Rightarrow K + a = K$. Thus the kernel of $g$ is $K$, the zero element of $R/K$. Therefore, $g$ is one-to-one and $R/K \cong S$.

**Theorem E.** If $f : R \rightarrow S$ is a ring homomorphism from $R$ onto $S$, and $A$ is an ideal of $R$, then $f(A)$ is an ideal of $S$.

**Proof:** Suppose $c, d \in f(A)$. Then there exist $a, b$ in $A$ such that $f(a) = c$ and $f(b) = d$. Now $A$ is an ideal so that $a - b \in A \Rightarrow f(a - b) = c - d \in f(A)$. Also, if $e \in S$, there exists $r \in R$ such that $f(r) = e$. Then $ar \in A \Rightarrow f(ar) = c e \in f(A)$. Note that $A$ nonempty $\Rightarrow f(A)$ is nonempty. Thus $f(A)$ is an ideal.

**Theorem F.** If $f : R \rightarrow S$ is a ring homomorphism from $R$ onto
S, and \( B \) is an ideal of \( S \), then \( f^{-1}(B) \) is an ideal of \( R \) and kernel \( f = f^{-1}(B) \).

**Proof:** Let \( a, b \in f^{-1}(B) \Rightarrow \) there exist \( s, t \in B \) such that \( f(a) = s, f(b) = t \). Now \( B \) is an ideal so that \( s - t \in B \). Since \( f(a - b) = s - t \), we have \( a - b \in f^{-1}(B) \). Also, if \( r \in R, f(r) = v \in S \). \( B \) is an ideal so that \( sv \in B \). Now \( f(ar) = sv \Rightarrow ar \in f^{-1}(B) \). Again if \( B \neq \emptyset \), \( f^{-1}(B) \) is non-empty so \( f^{-1}(B) \) is an ideal. Note that \( 0 \in B \) for any ideal \( B \) so kernel \( f = f^{-1}(B) \).

**Corollary P.1** There is a one-to-one correspondence between the ideals of \( S \) and those of \( R \) which contain \( K \), the kernel of \( f \).

**Proof:** From the previous theorem we know that for any ideal of \( S \) there is an ideal of \( R \) containing \( K \). We must show there is precisely one. Suppose \( A \) and \( B \) are ideals of \( R \) such that \( K \subseteq A, K \subseteq B, \) and \( f(A) = f(B) = C \), some ideal of \( S \). Let \( a \in A \); then \( f(a) = c \) and there exists \( b \in B \) such that \( f(b) = c \). Therefore \( f(a - b) = 0 \). Hence \( a - b \in K \Rightarrow a - b \in B \) since \( K \subseteq B \). Therefore \( b + (a - b) = a \in B \). Similarly, if \( b \in B \), then \( b \in A \). Therefore \( A = B \) and there is precisely one ideal of \( R \) containing \( K \) such that its image is \( C \).

**Theorem G.** Let \( R \) be a commutative ring, and \( S \) a nonempty subset which is closed under multiplication. Let \( A \) be an ideal of \( R \) with \( S \cap A \) empty. Then there is a prime ideal, \( P \), of \( R \) with \( A \subseteq P \) and \( P \cap S \) empty.
Proof: Since the proof is based directly on Zorn's Lemma, we state that now: If $L$ is a nonempty partially ordered set such that every linearly ordered subset of $L$ has an upper bound in $L$, then $L$ has maximal elements. Now let $C = \{D | D$ is an ideal of $R$, $A \subseteq D$, and $D \cap S$ is empty $\}$. Note $C$ contains $A$ and is therefore nonempty. Also $C$ is partially ordered by set inclusion and every linearly ordered subset $B$ of $C$ has an upper bound in $C$, namely $\bigcup B$. So $C$ has maximal elements. Let $E$ be a maximal element of $C$. We claim $E$ is prime for, if not, there exists $a \in E$ such that $a \notin E$ and $c \notin E$. But then $E \subset E + (a)$ and $E \subset E + (c)$. Thus \((E + (a)) \cap S\) must be nonempty and \((E + (c)) \cap S\) is nonempty since $E$ was maximal in $C$. But if $s_1 = e_1 + r_1 a \in S$ and $s_2 = e_2 + r_2 c \in S$, we get $s_1 s_2 \in E \Rightarrow E \cap S \neq \emptyset$. This contradiction concludes the proof; i.e., $E$ is prime.

Corollary G.1 Let $R$ be a ring with 1, $A$ an ideal in $R$, $1 \notin A$. Then there is a prime ideal $P$ such that $A \subseteq P$ and $1 \notin P$.

Proof: This follows directly from Theorem G since $S = \{1\}$ is multiplicatively closed.

Corollary G.2 If $R$ is a commutative ring, then $\sqrt{(0)} = \cap \{P | P$ is a prime ideal of $R\} = \{x \in R | x^n = 0$ for some $n\}$.

Proof: If $x^n = 0$, then $x^n \in P \Rightarrow x \in P$ for all prime ideals. If $x^n \neq 0$ for all $n$, then $\{x^n | n = 1, 2, \ldots\} = S$ is multiplicatively closed. Also $\{0\} \cap S$ is empty so there exists
a prime ideal \( P \) such that \( \{0\} \subseteq P \) and \( S \cap P \) is empty. But
\( S \cap P \) empty \( \Rightarrow x \notin P \Rightarrow x \notin \cap \{P|P \text{ is a prime ideal of } R\} \).

**Corollary G.3** If \( R \) is a commutative ring, \( A \) is an ideal
in \( R \), then \( \sqrt{A} = \{x|x^n \in A \text{ for some } n\} \).

**Proof:** If \( x^n \in A \) and \( A \subseteq P \), then \( x^n \in P \) and \( x \in P \) so
\( x \in \sqrt{A} \). If \( x^n \notin A \) for all \( n \), then \( S \cap A = \emptyset \), where \( S \\
again equals \( \{x^n|n = 1, 2, \ldots\} \). Therefore there exists
a prime ideal \( P \) such that \( A \subseteq P \) and \( P \cap S = \emptyset \); i.e.,
\( x \notin P \Rightarrow x \notin \sqrt{A} \).

**Corollary G.4** If \( \bar{b} \in U(R/(n)) \) and \( x \in \sqrt{(n)} \), then
\( \bar{b} + \bar{x} \in U(R/(n)) \).

**Proof:** Consider \( \bar{b}^{-1}(\bar{b} + \bar{x}) = 1 + \bar{b}^{-1}x \). Now \( \bar{b}^{-1}x = b^{-1}x + (n) \). Since \( x \in \sqrt{(n)} \), \( \bar{b}^{-1}x \in \sqrt{(n)} \). Let \( \bar{b}^{-1}x = z \).

Then \( z \in \sqrt{(n)} \Rightarrow z^m \in (n) \) for some \( m \). Therefore \( \bar{z}^m = \bar{0} \).

Now let \( \bar{a} = \sum_{k=0}^{m-1} \bar{z}^k \). Then \( (1 - \bar{z})\bar{a} = \sum_{k=0}^{m-1} \bar{z}^k - \frac{m-1}{m} \bar{z}^{k+1} \)
\( \Rightarrow (\bar{1} + \bar{b} \bar{x})\bar{a} = \bar{1} - \bar{z}^m = \bar{1} \). Therefore, \( \bar{a} + \bar{a} \bar{b} \bar{x} = \bar{1} \)
or \( \bar{a} \bar{b}^{-1}(\bar{b} + \bar{x}) = \bar{1} \). Therefore \( \bar{b} + \bar{x} \) is a unit of \( R/(n) \).
REFERENCES


Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.