NUMBER THEORY AND LINEAR ALGEBRA: MM6B12

Shyama M.P.
Assistant Professor,
Department of Mathematics
Malabar Christian College, Calicut

February 3, 2014

Contents

MODULE I

Chapter 1

DIVISIBILITY THEORY IN THE INTEGERS

Well- Ordering Principle
Every non empty set S of nonnegative integers contains a least element. That is, there exists some integer a in S such that $a \leq b$ for all b in S.

THE DIVISION ALGORITHM
Division Algorithm, the result is familiar to most of us roughly, it asserts that an integer a can be "divided" by a positive integer b in such a way that the remainder is smaller than b. The exact statement of this fact is Theorem 1.:

Theorem 1. *Given integers a and b, with $b > 0$, there exist unique integers q and r satisfying*

$$a = qb + r \qquad 0 \leq r < b$$

The integers q and r are called, respectively, the quotient and remainder in the division of a by b.

Proof. Let a and b be integers with $b > 0$ and consider the set

$$S = \{a - xb : xisaninteger; a - xb \geq 0\}.$$

Claim: The set S is nonempty

It suffices to find a value x which making $a - xb$ nonnegative. Since $b \geq 1$, we have $|a|b \geq |a|$ and so, $a - (-|a|)b = a + |a|b \geq a + |a| \geq 0$. For the choice $x = -|a|$, then $a - xb$ lies in S. Therefore S is nonempty, hence the claim. Therefore by Well-Ordering Principle, S contains a small integer, say r. By

the definition of S there exists an integer q satisfying

$$r = a - qb \qquad 0 \leq r.$$

Claim: $r < b$

Suppose $r \geq b$. Then we have

$$a - (q+1)b = (a - qb) - b = r - b \geq 0.$$

This implies that, $a - (q+1)b \in S$. But $a - (q+1)b = r - b < r$, since $b > 0$, leading to a contradiction of the choice of r as the smallest member of S. Hence, $r < b$, hence the claim.
Next we have to show that the uniqueness of q and r. Suppose that a as two representations of the desired form, say,

$$a = qb + r = q'b + r',$$

where $0 \leq r < b$ and $0 \leq r' < b$. Then $(r' - r) = b(q - q')$. Taking modulus on both sides,

$$|(r' - r)| = |b(q - q')| = |b||(q - q')| = b|(q - q')|.$$

But we have $-b < -r \leq 0$ and $0 \leq r' < b$, upon adding these inequalities we obtain $-b < r' - r < b$. This implies $b|(q - q')| < b$, which yields $0 \leq |q - q'| < 1$. Because $|q - q'|$ is a nonnegative integer, the only possibility is that $|q - q'| = 0$, hence, $q = q'$. This implies $|r' - r| = 0$, that is, $r = r'$. Hence the proof. □

Corollary 1. *If a and b are integers, with $b \neq 0$, then there exists integers q and r such that*

$$a = qb + r \qquad 0 \leq r < |b|.$$

Proof. It is enough to consider the case in which b is negative. Then $|b| > 0$, and Theorem 1. produces unique integers q' and r for which

$$a = q'|b| + r \qquad 0 \leq r < |b|.$$

Noting that $|b| = -b$, we may take $q = -q'$ to arrive at $a = qb + r$, with $0 \leq r < |b|$. □

Application of the Division Algorithm

1. Square of any integer is either of the form $4k$ or $4k+1$. That is, the square of integer leaves the remainder 0 or 1 upon division by 4.
   Solution: Let a be any integer. If a is even, we can let $a = 2n, n$ is an integer, then $a^2 = (2n)^2 = 4n^2 = 4k$. If a is odd, we can let $a = 2n+1, n$ is an integer, then $a^2 = (2n+1)^2 = 4n^2+4n+1 = 4(n^2+n)+1 = 4k+1$.

2. The square of any odd integer is of the form $8k+1$.
   Solution: Let a be an integer and let $b = 4$, then by division algorithm a is representable as one of the four forms: $4q, 4q+1, 4q+2, 4q+3$. In this representation, only those integers of the forms $4q+1$ and $4q+3$ are odd. If $a = 4q+1$, then

   $$a^2 = (4q+1)^2 = 16q^2 + 8q + 1 = 8(2q^2+q) + 1 = 8k+1.$$

   If $a = 4q+3$, then

   $$a^2 = (4q+3)^2 = 16q^2+24q+9 = 16q^2+24q+8+1 = 8(2q^2+3q+1)+1 = 8k+1.$$

3. For all integer $a \geq 1$, $\frac{a(a^2+2)}{3}$ is an integer.
   Solution: Let $a \geq 1$ be an integer. According to division algorithm, a is of the form $3q, 3q+1$ or $3q+2$. If $a = 3q$, then

   $$\frac{3q((3q)^2+2)}{3} = 9q^3 + 2q,$$

   which is clearly an integer. Similarly we can prove other two cases also.

THE GREATEST COMMON DIVISOR

Definition 1. *An integer b is said to be divisible by an integer $a \neq 0$, in symbols $a|b$, if there exists some integer c such that $b = ac$. We write $a \nmid b$ to indicate that b is not divisible by a.*

Thus, for example, -22 is divisible by 11, because $-22 = 11(-2)$. However, 22 is not divisible by 3; for there is no integer c that makes the statement $22 = 3c$ true.
There is other language for expressing the divisibility relation $a|b$. We could say that a is a divisor of b, that a is a factor of b, or that b is a multiple of a. Notice that in Definition 1 there is a restriction on the divisor a: Whenever the notation $a|b$ is employed, it is understood that a is different from zero.
If a is a divisor of b, then b is also divisible by $-a$ (indeed, $b = ac$ implies that $b = (-a)(-c)$), so that the divisors of an integer always occur in pairs.

To find all the divisors of a given integer, it is sufficient to obtain the positive divisors and then adjoin to them the corresponding negative integers. For this reason, we shall usually limit ourselves to a consideration of positive divisors. It will be helpful to list some immediate consequences of Definition 1.

Theorem 2. *For integers a, b, c, the following hold:*

1. *$a|0$, $1|a$, $a|a$.*
2. *$a|1$ if and only if $a = \pm 1$.*
3. *If $a|b$ and $c|d$, then $ac|bd$.*
4. *If $a|b$ and $b|c$, then $a|c$.*
5. *$a|b$ and $b|a$ if and only if $a = \pm b$.*
6. *If $a|b$ and $b \neq 0$, then $|a| \leq |b|$.*
7. *If $a|b$ and $a|c$, then $a|(bx + cy)$ for arbitrary integers x and y.*

Proof. 1. Since $0 = a.0$, $a|0$. Since $a = 1.a$, $1|a$. Since $a = a.1$, $a|a$.

2. We have $a|1$ if and only if $1 = a.c$ for some c, this is if and only if $a = \pm 1$.

3. Clear from definition.

4. Clear from definition.

5. Clear from definition.

6. If $a|b$, then there exists an integer c such that $b = ac$; also, $b \neq 0$ implies that $c \neq 0$. Upon taking absolute values, we get $|b| = |ac| = |a||c|$. Because $c \neq 0$, it follows that $|c| \geq 1$, whence $|b| = |a||c| \geq |a|$.

7. The relations $a|b$ and $a|c$ ensure that $b = ar$ and $c = as$ for suitable integers r and s. But then whatever the choice of x and y, $bx + cy = arx + asy = a(rx + sy)$. Because $rx + sy$ is an integer, this says that $a|(bx + cy)$, as desired.

□

Definition 2. *Let a and b be given integers, with at least one of them different from zero. The greatest common divisor of a and b, denoted by $gcd(a, b)$, is the positive integer d satisfying the following:*

(i) *$d|a$ and $d|b$.*

(ii) *If $c|a$ and $c|b$, then $c \leq d$.*

Example: The positive divisors of -12 are $1, 2, 3, 4, 6, 12$, whereas those of 30 are $1, 2, 3, 5, 6, 10, 15, 30$; hence, the positive common divisors of -12 and 30 are $1, 2, 3, 6$. Because 6 is the largest of these integers, it follows that $gcd(-12, 30) = 6$. In the same way, we can show that $gcd(-5, 5) = 5$, $gcd(8, 17) = 1$, $gcd(-8, -36) = 4$.

Theorem 3. *Given integers a and b, not both of which are zero, there exist integers x and y such that*

$$gcd(a, b) = ax + by.$$

Proof. Consider the set S of all positive linear combinations of a and b :

$$S = \{au + bv : au + bv > 0; u, v \; integers\}.$$

Since, if $a \neq 0$ then $|a| = au+b.0 \in S$, where $u = 1, if a > 0; u = -1, if a < 0$, S is nonempty. Therefore by the Well-Ordering Principle, S must contain a smallest element, say d. Thus, from the very definition of S, there exist integers x and y for which $d = ax + by$.
Claim: $d = gcd(a, b)$

By using the Division Algorithm, we can obtain integers q and r such that $a = qd + r$, where $0 \leq r < d$. Then r can be written in the form:

$$\begin{aligned} r &= a - qd \\ &= a - q(ax + by) \\ &= a(1 - qx) + b(-qy). \end{aligned}$$

If r were positive, then this representation would imply that r is a member of S, contradicting the fact that d is the least integer in S (recall that $r < d$). Therefore, $r = 0$, and so $a = qd$, or equivalently $d|a$. By similar reasoning, $d|b$, this implies d is a common divisor of a and b.
Now if c is an arbitrary positive common divisor of the integers a and b, then part (7) of Theorem 2 allows us to conclude that $c|(ax + by)$; that is, $c|d$. By part (6) of the same theorem, $c = |c| \leq |d| = d$, so that d is greater than every positive common divisor of a and b. Hence $d = gcd(a, b)$. Hence the claim. Therefore $gcd(a, b) = ax + by$. □

Corollary 2. *If a and b are given integers, not both zero, then the set*

$$T = ax + by : x, y \; are \; integers$$

is precisely the set of all multiples of $d = gcd(a, b)$.

Proof. Because $d|a$ and $d|b$, we know that $d|(ax + by)$ for all integers x, y. Thus, every member of T is a multiple of d. Conversely, d may be written as $d = ax_0 + by_0$ for suitable integers x_0 and y_0 , so that any multiple nd of d is of the form

$$nd = n(ax_0 + by_0) = a(nx_0) + b(ny_0).$$

Hence, nd is a linear combination of a and b, and, by definition, lies in T. □

Definition 3. *Two integers a and b, not both of which are zero, are said to be relatively prime whenever $gcd(a, b) = 1$.*

Theorem 4. *Let a and b be integers, not both zero. Then a and b are relatively prime if and only if there exist integers x and y such that $1 = ax + by$.*

Proof. If a and b are relatively prime so that $gcd(a, b) = 1$, then Theorem 3 guarantees the existence of integers x and y satisfying $1 = ax+by$. Conversely, suppose that $1 = ax + by$ for some choice of x and y, and that $d = gcd(a, b)$. Because $d|a$ and $d|b$, Theorem 2 yields $d|(ax+by)$, or $d|1$. This implies $d = \pm 1$. But d is a positive integer, $d = 1$. That is a and b are relatively prime. □

Corollary 3. *If $gcd(a, b) = d$, then $gcd(a/d, b/d) = 1$.*

Proof. Since $d|a$ and $d|b$, a/d and b/d are integers. We have, if $gcd(a, b) = d$, then there exists x and y such that $d = ax + by$. Upon dividing each side of this equation by d, we obtain the expression

$$1 = (a/d)x + (b/d)y.$$

Because a/d and b/d are integers, a/d and b/d are relatively prime. Therefore $gcd(a/d, b/d) = 1$. □

Corollary 4. *If $a|c$ and $b|c$, with $gcd(a, b) = 1$, then $ab|c$.*

Proof. Since $a|c$ and $b|c$, we can find integers r and s such that $c = ar = bs$. Given that $gcd(a, b) = 1$, so there exists integers x and y such that $1 = ax+by$. Multiplying the last equation by c, we get,

$$c = c1 = c(ax + by) = acx + bcy.$$

If the appropriate substitutions are now made on the right-hand side, then

$$c = a(bs)x + b(ar)y = ab(sx + ry).$$

This implies, $ab|c$. □

Theorem 5. *(Euclid's lemma.) If $a|bc$, with $gcd(a,b) = 1$, then $a|c$.*

Proof. Since $gcd(a,b) = 1$, we have $1 = ax + by$ for some integers x and y. Multiplication of this equation by c produces

$$c = 1c = (ax + by)c = acx + bcy.$$

Since $a|bc$ and $a|ac$, we have $a|acx + bcy$. This implies $a|c$. □

Note: If a and b are not relatively prime, then the conclusion of Euclid's lemma may fail to hold. For example: $6|9.4$ but $6 \nmid 9$ and $6 \nmid 4$.

Theorem 6. *Let a, b be integers, not both zero. For a positive integer d, $d = gcd(a,b)$ if and only if*

(i) *$d|a$ and $d|b$.*

(ii) *Whenever $c|a$ and $c|b$, then $c|d$.*

Proof. Suppose that $d = gcd(a,b)$. Certainly, $d|a$ and $d|b$, so that (i) holds. By Theorem 3, d is expressible as $d = ax + by$ for some integers x, y. Thus, if $c|a$ and $c|b$, then $c|(ax + by)$, or rather $c|d$. This implies, condition (ii) holds. Conversely, let d be any positive integer satisfying the stated conditions (i) and (ii). Given any common divisor c of a and b, we have $c|d$ from hypothesis (ii). This implies that $d \geq c$, and consequently d is the greatest common divisor of a and b. □

THE EUCLIDEAN ALGORITHM

Lemma 1. *If $a = qb + r$, then $gcd(a,b) = gcd(b,r)$.*

Proof. If $d = gcd(a,b)$, then the relations $d|a$ and $d|b$ together imply that $d|(a - qb)$, or $d|r$. Thus, d is a common divisor of both b and r. On the other hand, if c is an arbitrary common divisor of b and r, then $c|(qb + r)$, whence $c|a$. This makes c a common divisor of a and b, so that $c \leq d$. It now follows from the definition of $gcd(b,r)$ that $d = gcd(b,r)$. □

The Euclidean algorithm

The Euclidean Algorithm may be described as follows: Let a and b be two integers whose greatest common divisor is desired. Because $gcd(|a|,|b|) = gcd(a,b)$,, with out loss of generality we may assume $a \geq b > 0$. The first step is to apply the Division Algorithm to a and b to get

$$a = q_1 b + r_1 \qquad 0 \leq r_1 < b.$$

If it happens that $r_1 = 0$, then $b|a$ and $gcd(a, b) = b$. When $r_1 \neq 0$, divide b by r_1 to produce integers q_2 and r_2 satisfying

$$b = q_2 r_1 + r_2 \qquad 0 \leq r_2 < r_1.$$

If $r_2 = 0$, then we stop; otherwise, proceed as before to obtain

$$r_1 = q_3 r_2 + r_3 \qquad 0 \leq r_3 < r_2.$$

This division process continues until some zero remainder appears, say, at the $(n+l)^{th}$ stage where r_{n-1} is divided by r_n (a zero remainder occurs sooner or later because the decreasing sequence $b > r_1 > r_2 > \cdots \geq 0$ cannot contain more than b integers). The result is the following system of equations:

$$\begin{aligned}
a &= q_1 b + r_1 \qquad 0 \leq r_1 < b \\
b &= q_2 r_1 + r_2 \qquad 0 \leq r_2 < r_1 \\
r_1 &= q_3 r_2 + r_3 \qquad 0 \leq r_3 < r_2 \\
&\vdots \\
r_{n-2} &= q_n r_{n-1} + r_n \qquad 0 \leq r_n < r_{n-1} \\
r_{n-1} &= q_{n+1} r_n + 0.
\end{aligned}$$

By Lemma 1,

$$gcd(a, b) = gcd(b, r_1) = \ = gcd(r_{n-1}, r_n) = gcd(r_n, 0) = r_n.$$

Note: Start with the next-to-last equation arising from the Euclidean Algorithm, we can determine x and y such that $gcd(a, b) = ax + by$.
Example: Let us see how the Euclidean Algorithm works in a concrete case by calculating, say, $gcd(12378, 3054)$. The appropriate applications of the Division Algorithm produce the equations

$$\begin{aligned}
12378 &= 4.3054 + 162 \\
3054 &= 18.162 + 138 \\
162 &= 1.138 + 24 \\
138 &= 5.24 + 18 \\
24 &= 1.18 + 6 \\
18 &= 3.6 + 0
\end{aligned}$$

This tells us that the last nonzero remainder appearing in these equations, namely, the integer 6, is the greatest common divisor of 12378 and 3054:

$$6 = gcd(12378, 3054).$$

To represent 6 as a linear combination of the integers 12378 and 3054, we start with the next-to-last of the displayed equations and successively eliminate the remainders 18, 24, 138, and 162:

$$\begin{aligned} 6 &= 24 - 18 \\ &= 24 - (138 - 5.24) \\ &= 6.24 - 138 \\ &= 6(162 - 138) - 138 \\ &= 6.162 - 7.138 \\ &= 6.162 - 7(3054 - 18.162) \\ &= 132.162 - 7.3054 \\ &= 132(12378 - 4.3054) - 7.3054 \\ &= 132.12378 + (-535)3054 \end{aligned}$$

Thus, we have

$$6 = gcd(12378, 3054) = 12378x + 3054y,$$

where $x = 132$ and $y = -535$. Note that this is not the only way to express the integer 6 as a linear combination of 12378 and 3054; among other possibilities, we could add and subtract 3054.12378 to get

$$\begin{aligned} 6 &= (132 + 3054)12378 + (-535 - 12378)3054 \\ &= 3186.12378 + (-12913)3054. \end{aligned}$$

Theorem 7. *If $k > 0$, then $gcd(ka, kb) = k\ gcd(a, b)$.*

Proof. If each of the equations appearing in the Euclidean Algorithm for a and b, multiplied by k, we obtain

$$\begin{aligned} ak &= q_1(bk) + r_1k \qquad 0 \le r_1k < bk \\ bk &= q_2(r_1k) + r_2k \qquad 0 \le r_2k < r_1k \\ &\vdots \\ r_{n-2}k &= q_n(r_{n-1}k) + r_nk \qquad 0 \le r_nk < r_{n-1}k \\ r_{n-1}k &= q_{n+1}(r_nk) + 0. \end{aligned}$$

But this is clearly the Euclidean Algorithm applied to the integers ak and bk, so that their greatest common divisor is the last nonzero remainder $r_n k$; that is,

$$gcd(ka, kb) = r_n k = k\ gcd(a, b),$$

Hence the theorem. □

Corollary 5. *For any integer $k \neq 0$, $gcd(ka, kb) = |k|\ gcd(a, b)$.*

Proof. We already have, if $k > 0$, then $gcd(ka, kb) = k\ gcd(a, b)$. Therefore it suffices to consider the case in which $k < 0$. Then $-k = |k| > 0$ and, by Theorem 7,

$$\begin{aligned} gcd(ak, bk) &= gcd(-ak, -bk) \\ &= gcd(a|k|, b|k|) \\ &= |k|\ gcd(a, b). \end{aligned}$$

Hence the result. □

Definition 4. *The least common multiple of two nonzero integers a and b, denoted by $lcm(a, b)$, is the positive integer m satisfying the following:*

(i) *$a|m$ and $b|m$.*

(ii) *If $a|c$ and $b|c$, with $c > 0$, then $m \leq c$.*

As an example, the positive common multiples of the integers -12 and 30 are 60, 120, 180, ... hence, $lcm(-12, 30) = 60$.

Theorem 8. *For positive integers a and b*

$$gcd(a, b)\ lcm(a, b) = ab.$$

Proof. Let $d = gcd(a, b)$ and let $m = ab/d$, then $m > 0$.
Claim: $m = lcm(a, b)$

Since d is the common divisor of a and b we have $a = dr$, $b = ds$ for integers r and s. Then $m = as = rb$. This implies, m a (positive) common multiple of a and b.
Now let c be any positive integer that is a common multiple of a and b, then $c = au = bv$ for some integers u and v. As we know, there exist integers x and y satisfying d =ax+ by. In consequence,

$$\frac{c}{m} = \frac{cd}{ab} = \frac{c(ax + by)}{ab} = (\frac{c}{b})x + \frac{c}{a})y = vx + uy.$$

This equation states that $m|c$, this implies, $m \leq c$. By the definition of least common multiple, we have $m = lcm(a, b)$. Hence the claim. Therefore $gcd(a, b)\ lcm(a, b) = ab$. □

Corollary 6. *For any choice of positive integers a and b, $lcm(a,b) = ab$ if and only if $gcd(a,b) = 1$.*

Definition 5. *If a, b, c, are three integers, not all zero, $gcd(a,b,c)$ is defined to be the positive integer d having the following properties:*

(i) *d is a divisor of each of a, b, c.*

(ii) *If e divides the integers a, b, c, then $e \leq d$.*

For example $gcd(39, 42, 54) = 3$ and $gcd(49, 210, 350) = 7$.

THE DIOPHANTINE EQUATION $ax + by = c$

The simplest type of Diophantine equation that we shall consider is the linear Diophantine equation in two unknowns:

$$ax + by = c,$$

where a, b, c are given integers and a, b are not both zero. A solution of this equation is a pair of integers $x_0,\ y_0$ that, when substituted into the equation, satisfy it; that is, we ask that $ax_0 + by_0 = c$.

Theorem 9. *The linear Diophantine equation $ax + by = c$ has a solution if and only if $d|c$, where $d = gcd(a,b)$. If $x_0,\ y_0$ is any particular solution of this equation, then all other solutions are given by*

$$x = x_0 + (\frac{b}{d})t \qquad y = y_0 - (\frac{a}{d})t,$$

where t is an arbitrary integer.

Example: Consider the linear Diophantine equation

$$172x + 20y = 1000$$

Applying the Euclidean's Algorithm to the evaluation of $gcd(172, 20)$, we find that

$$\begin{aligned} 172 &= 8.20 + 12 \\ 20 &= 1.12 + 8 \\ 12 &= 1.8 + 4 \\ 8 &= 24, \end{aligned}$$

whence $gcd(172, 20) = 4$. Because $4|1000$, a solution to this equation exists. To obtain the integer 4 as a linear combination of 172 and 20, we work

backward through the previous calculations, as follows:

$$\begin{aligned} 4 &= 12 - 8 \\ &= 12 - (20 - 12) \\ &= 212 - 20 \\ &= 2(172 - 8.20) - 20 \\ &= 2.172 + (-17)20 \end{aligned}$$

Upon multiplying this relation by 250, we arrive at

$$\begin{aligned} 1000 &= 250.4 \\ &= 250(2.172 + (-17)20) \\ &= 500.172 + (-4250)20, \end{aligned}$$

so that $x = 500$ and $y = -4250$ provide one solution to the Diophantine equation in question. All other solutions are expressed by

$$x = 500 + (20/4)t = 500 + 5t$$

$$y = -4250 - (172/4)t = -4250 - 43t,$$

for some integer t.
If we want to find positive solution,if any happen to exist. For this, t must be chosen to satisfy simultaneously the inequalities

$$5t + 500 > 0 \qquad -43t - 4250 > 0$$

or

$$-98\frac{36}{43} > t > -100.$$

Because t must be an integer, we are forced to conclude that $t = -99$. Thus, our Diophantine equation has a unique positive solution $x = 5$, $y = 7$ corresponding to the value $t = -99$.

Corollary 7. *If $gcd(a, b) = 1$ and if x_0 , y_0 is a particular solution of the linear Diophantine equation $ax + by = c$, then all solutions are given by*

$$x = x_0 + bt \qquad y = Y_0 - at,$$

for integral values of t.

Problem 1: A customer bought a dozen pieces of fruit, apples and oranges, for | 1.32. If an apple costs 3 rupees more than an orange and more apples than oranges were purchased, how many pieces of each kind were bought?

Solution: Let x be the number of apples and y be the number of oranges purchased. Also let z be the cost (in rupees) of an orange. Then the conditions of the problem lead to:

$$(z+3)x + zy = 132,$$

or equivalently

$$3x + (x+y)z = 132.$$

Because $x + y = 12$, the previous equation may be replaced by

$$3x + 12z = 132$$

this implies,

$$x + 4z = 44.$$

Now the problem is to find integers x and z satisfying the Diophantine equation $x + 4z = 44$. We have $gcd(1, 4) = 1$ is a divisor of 44, there is a solution to this equation. Upon multiplying the relation 1 = 1 (-3) + 4 1 by 44 to get

$$44 = 1(-132) + 4.44 \tag{1.1}$$

it follows that $x_0 = -132$, $z_0 = 44$, serves as one solution. All other solutions of Eq. (1.1) are of the form

$$x = -132 + 4t \qquad z = 44 - t,$$

where t is an integer.
Not all of the choices fort furnish solutions to the original problem. Only values of t that ensure $12 \geq x > 6$ should be considered. This requires obtaining those values of t such that

$$12 \geq -132 + 4t > 6.$$

Now,$12 \geq -132 + 4t$ implies that $t \leq 36$, whereas $-132 + 4t > 6$ gives $t > 34\frac{1}{2}$. The only integral values of t to satisfy both inequalities are $t = 35$ and $t = 36$. Thus, there are two possible purchases: a dozen apples costing 11 rupees apiece (the case where $t = 36$), or 8 apples at 12 rupees each and 4 oranges at 9 rupees each (the case where $t = 35$).
Problem 2: If a cock is worth 5 coins, a hen 3 coins, and three chicks together 1 coin, how many cocks, hens, and chicks, totaling 100, can be bought for 100 coins?
Solution: Let x be the number of cocks, y be the number of hens, z be the number of chicks. Then the conditions of the problem lead to

$$5x + 3y + \frac{1}{3}z = 100 \qquad x + y + z + 100.$$

Eliminating one of the unknowns, we are left with a linear Diophantine equation in the two other unknowns. Specifically, because the quantity $z = 100 - x - y$, we have $5x + 3y + \frac{1}{3}(100 - x - y) = 100$, or

$$7x + 4y = 100.$$

This equation has the general solution $x = 4t$, $y = 25 - 7t$, so that $z = 75 + 3t$, where t is an arbitrary integer. Some solutions are:

$$\begin{array}{cccccccc} x & = & 4 & y & = & 18 & z & = & 78 \\ x & = & 8 & y & = & 11 & z & = & 81 \\ x & = & 12 & y & = & 4 & z & = & 84 \end{array}$$

To obtain all solutions in the positive integers, t must be chosen to satisfy simultaneously the inequalities

$$4t > 0 \quad 25 - 7t > 0 \quad 75 + 3t > 0.$$

The last two of these are equivalent to the requirement $-25 < t < 3\frac{4}{7}$. Because t must have a positive value, we conclude that $t = 1, 2, 3$, leading to precisely the values given above.

Chapter 2

PRIMES AND THEIR DISTRIBUTION

THE FUNDAMENTAL THEOREM OF ARITHMETIC

Definition 6. *An integer $p > 1$ is called a prime number, or simply a prime, if its only positive divisors are 1 and p. An integer greater than 1 that is not a prime is termed composite.*

Among the first ten positive integers, 2, 3, 5, 7 are primes and 4, 6, 8, 9, 10 are composite numbers. Note that the integer 2 is the only even prime, and according to our definition the integer 1 plays a special role, being neither prime nor composite.

Theorem 10. *If p is a prime and $p|ab$, then $p|a$ or $p|b$.*

Proof. If $p|a$, then we need go no further, so let us assume that $p \nmid a$. Because the only positive divisors of p are 1 and p itself, this implies that $gcd(p, a) = 1$. Hence, by Euclid's lemma, we get $p|b$. □

Corollary 8. *If p is a prime and $p|a_1a_2 \cdots a_n$, then $p|a_k$ for some k, where $1 \leq k \leq n$.*

Proof. We proceed by induction on n, the number of factors. When $n = 1$, the stated conclusion obviously holds; whereas when $n = 2$, the result is the content of Theorem 10. Suppose, as the induction hypothesis, that $n > 2$ and that whenever p divides a product of less than n factors, it divides at least one of the factors. Now let $p|a_1a_2 \cdots a_n$. From Theorem 10, either $p|a_n$ or $p|a_1a_2 \cdots a_{n-1}$ If $p|a_n$, then we are through. As regards the case where $p|a_1a_2 \cdots a_{n-1}$, the induction hypothesis ensures that $p|a_k$ for some choice of k, with $1 \leq k \leq n - 1$. In any event, p divides one of the integers $a_1, a_2, \cdots, a_n$. □

Corollary 9. *If $p,\ q_1,\ q_2, \cdots,\ q_n$ are all primes and $p|q_1q_2\cdots q_n$, then $p = q_k$ for some k, where $1 \leq k \leq n$.*

Proof. By Corollary 8, we know that $p|q_k$ for some k, with $1 \leq k \leq n$. Being a prime, q_k is not divisible by any positive integer other than 1 or q_k itself. Because $p > 1$, we are forced to conclude that $p = q_k$. □

Theorem 11. *(Fundamental Theorem of Arithmetic.) Every positive integer $n > 1$ can be expressed as a product of primes; this representation is unique, apart from the order in which the factors occur.*

Proof. Either n is a prime, there is nothing to prove. If n is composite, then there exists an integer d satisfying $d|n$ and $1 < d < n$. Among all such integers d, choose p_1 to be the smallest (this is possible by the Well-Ordering Principle). Then P_1 must be a prime number. Otherwise it too would have a divisor q with $1 < q < p_1$; but then $q|p_1$ and $p_1|n$ imply that $q|n$, which contradicts the choice of p_1 as the smallest positive divisor, not equal to 1, of n. We therefore may write $n = p_1n_1$, where p_1 is prime and $1 < n_1 < n$.If n_1 happens to be a prime, then we have our representation. In the contrary case, the argument is repeated to produce a second prime number p_2 such that $n_1 = p_2n_2$; that is,

$$n = p_1P_2n_2 \qquad 1 < n_2 < n_1.$$

If n_2 is a prime,then it is not necessary to go further. Otherwise, write $n_2 = p_3n_3$, with p_3 a prime:

$$n = p_1P_2p_3n_3 \qquad 1 < n_3 < n_2.$$

The decreasing sequence $n > n_1 > n_2 > \cdots > 1$ cannot continue indefinitely, so that after a finite number of steps n_{k-1} is a prime, call it, p_k. This leads to the prime factorization

$$n = p_1p_2\cdots p_k.$$

To establish the second part of the proof-the uniqueness of the prime factorization, let us suppose that the integer n can be represented as a product of primes in two ways, say,

$$n = p_1p_2\cdots p_r = q_1q_2\cdots q_s \qquad r \leq s,$$

where the p_i and q_j are all primes, written in increasing magnitude so that

$$p_1 \leq p_2 \leq \cdots \leq p_r \qquad q_1 \leq q_2 \leq \cdots \leq q_s.$$

Because $p_1|q_1q_2\cdots q_s$, Corollary 9 tells us that $p_1 = q_k$ for some k; but then $p_1 \geq q_1$. Similar reasoning gives $q_1 \geq p_1$, whence $p_1 = q_1$. We may cancel this common factor and obtain

$$p_2p_3\cdots p_r = q_2q_3\cdots q_s.$$

Now repeat the process to get $p_2 = q_2$ and, in turn,

$$p_3p_4\cdots p_r = q_3q_4\cdots q_s.$$

Continue in this fashion. If the inequality $r < s$ were to hold, we would eventually arrive at

$$1 = q_{r+1}q_{r+2}\cdots q_s,$$

which is absurd, because each $q_j > 1$. Hence, $r = s$ and

$$p_1 = q_1,\ p_2 = q_2,\ \cdots, p_r = q_r,$$

making the two factorizations of n identical. The proof is now complete. □

Corollary 10. *Any positive integer $n > 1$ can be written uniquely in a canonical form*

$$n = p_1^{k_1}p_2^{k_2}\cdots p_r^{k_r},$$

where, for $i = 1, 2, \cdots, r$, each k_i is a positive integer and each p_i is a prime, with $p_1 < p_2 < \cdots < p_r$.

Theorem 12. *(Pythagoras.) The number $\sqrt{2}$ is irrational.*

Proof. Suppose, to the contrary, that $\sqrt{2}$ is a rational number, say, $\sqrt{2} = a/b$, where a and b are both integers with $gcd(a, b) = 1$. Squaring, we get $a^2 = 2b^2$, so that $b|a^2$. Claim: $b = 1$ If $b > 1$, then the Fundamental Theorem of Arithmetic guarantees the existence of a prime p such that $p|b$. It follows that $p|a^2$. This implies that $p|a$; hence, $gcd(a, b) \geq p$. We therefore arrive at a contradiction, unless $b = 1$, hence the claim.
But if this happens, then $a^2 = 2$, which is impossible. Our supposition that $\sqrt{2}$ is a rational number is not true, and so $\sqrt{2}$ must be irrational. □

THE SIEVE OF ERATOSTHENES

Given a particular integer, how can we determine whether it is prime or composite and, in the latter case, how can we actually find a nontrivial divisor? The most obvious approach consists of successively dividing the integer in question by each of the numbers preceding it; if none of them (except 1) serves as a divisor, then the integer must be prime. Although this method is very simple to describe, it cannot be regarded as useful in practice. For even if one is undaunted by large calculations, the amount of time and work involved may be prohibitive.

Theorem 13. *(Euclid.) There is an infinite number of primes.*

Proof. Euclid's proof is by contradiction. Let $p_1 = 2,\ p_2 = 3,\ p_3 = 5,\ p_4 = 7, \cdots$ be the primes in ascending order, and suppose that there is a last prime, called p_n Now consider the positive integer $P = p_1 p_2 \cdots p_{n+1}$. Because $P > 1$, we have P is divisible by some prime p. But $p_1,\ p_2, \cdots, p_n$ are the only prime numbers, so that p must be equal to one of $p_1, p_2, ..., p_n$. Combining the divisibility relation $p|p_1, p_2, \cdots, p_n$ with $p|P$, we arrive at $p|P - p_1, p_2, \cdots, p_n$ or, equivalently, $p|1$. The only positive divisor of the integer 1 is 1 itself and, because $p > 1$, a contradiction arises. Thus, no finite list of primes is complete, whence the number of primes is infinite. □

Definition 7. *For a prime p, define $p^{\#}$ = the product of all primes that are less than or equal to p. Numbers of the form $p^{\#} + 1$ called Euclidean numbers.*

Note: Not all Euclidean numbers are primes. For example, $13^{\#} + 1 = 59.509$, $17^{\#} + 1 = 19.97.277$, $19^{\#} + 1 = 347.27953$.

Theorem 14. *If p_n is the nth prime number, then $p_n \leq 2^{2^{n-l}}$.*

Proof. Proof is by induction on n.When $n = 1$, $p_n = p_1 = 2$ and $2^{2^{n-1}} = 2^{2^{1-1}} = 2^{2^0} = 2^1 = 2$, the result is true.Suppose that $n > 1$ and that the result holds for all integers up to n. Then

$$\begin{array}{rcl} p_{n+1} & \leq & p_1 p_2 \cdots p_n + 1 \\ & \leq & 2 2^2 \cdots 2^{n-1} + 1 \quad = \quad 2^{1+2+2^2+\cdots+2^{n-1}} + 1. \end{array}$$

Recalling the identity $1 + 2 + 2^2 + \ + 2^{n-1} = 2^n - 1$, we obtain

$$p_{n+1} \leq 2^{2^n - 1} + 1.$$

However, $1 \leq 2^{2^n - 1}$ for all n; whence

$$\begin{array}{rcl} p_{n+1} & \leq & 2^{2^n - 1} + 2^{2^n - 1} \\ & = & 2.2^{2^n - 1} \quad = \quad 2^{2^n} \end{array}$$

completing the induction step, and the argument. □

Corollary 11. *For $n \geq 1$, there are at least $n + 1$ primes less than 2^{2^n}.*

Proof. From the theorem, we know that $p_1, p_2, \cdots, p_{n+1}$ are all less than 2^{2^n}. □

Chapter 3

THE THEORY OF CONGRUENCES

Definition 8. *Let n be a fixed positive integer. Two integers a and b are said to be congruent modulo n, symbolized by*

$$a \equiv b(mod n)$$

if n divides the difference $a - b$; that is, provided that $a - b = kn$ for some integer k.

Theorem 15. *For arbitrary integers a and b, $a \equiv b(mod n)$ if and only if a and b leave the same nonnegative remainder when divided by n.*

Proof. Suppose $a \equiv b(mod n)$, so that $a = b + kn$ for some integer k. Upon division by n, b leaves a certain remainder r; that is, $b = qn + r$, where $0 \leq r < n$. Therefore,

$$a = b + kn = (qn + r) + kn = (q + k)n + r$$

which indicates that a has the same remainder as b.
On the other hand, suppose we can write $a = q_1 n + r$ and $b = q_2 n + r$, with the same remainder r $(0 \leq r < n)$. Then

$$a - b = (q_1 n + r) - (q_2 n + r) = (q_1 - q_2)n,$$

whence $n|a - b$. That is, $a \equiv b(mod n)$. □

Theorem 16. *Let $n > 1$ be fixed and a, b, c, d be arbitrary integers. Then the following properties hold:*

1. $a \equiv a(mod n)$.

2. *If* $a \equiv b(modn)$, *then* $b \equiv a(modn)$.

3. *If* $a \equiv b(modn)$ *and* $b \equiv c(modn)$, *then* $a \equiv c(modn)$.

4. *If* $a \equiv b(modn)$ *and* $c \equiv d(modn)$, *then* $a + c \equiv b + d(modn)$ *and* $ac \equiv bd(modn)$.

5. *If* $a \equiv b(modn)$, *then* $a + c \equiv b + c(modn)$ *and* $ac \equiv bc(modn)$.

6. *If* $a \equiv b(modn)$, *then* $a^k \equiv b^k(modn)$ *for any positive integer* k.

Problem 1: Show that $41|2^{20} - 1$.
Solution: We have

$$2^5 \equiv -9(mod\ 41).$$

Therefore

$$(2^5)^4 \equiv (-9)^4(mod\ 41).$$

This implies that

$$2^{20} \equiv (-9)^4(mod\ 41).$$

But we have $(-9)^4 = 81.81$ and $81 \equiv -1(mod\ 41)$. Therefore

$$2^{20} \equiv (-1)(-1)(mod\ 41).$$

This implies $41|2^{20} - 1$.
Problem 2: Find the remainder obtained upon dividing the sum

$$1! + 2! + 3! + 4! + \cdots + 99! + 100!$$

by 12.
Solution: We have $4! \equiv 24 \equiv 0(mod\ 12)$; thus, for $k \geq 4$,

$$k! \equiv 4!.5.6 \cdots k \equiv 0.5..6 \cdots k \equiv 0(mod\ 12).$$

Therefore

$$1! + 2! + 3! + 4! + \cdots + 100! \equiv 1! + 2! + 3! + 0 + \cdots + 0 \equiv 9(mod\ 12).$$

The remainder 9.

Theorem 17. *If* $ca \equiv cb(mod\ n)$, *then* $a \equiv b(mod\ n/d)$, *where* $d = gcd(c, n)$.

Proof. By hypothesis, we can write

$$c(a-b) = ca - cb = kn, \tag{3.1}$$

for some integer k. Knowing that $gcd(c,n) = d$, there exist relatively prime integers r and s satisfying $c = dr$, $n = ds$. When these values are substituted in Eq. 3.1 and the common factor d canceled, the net result is

$$r(a-b) = ks.$$

Hence, $s|r(a-b)$ and $gcd(r,s) = 1$. Euclid's lemma yields $s|(a-b)$, which implies $a \equiv b(mod\ s)$; in other words, $a \equiv b(mod\ n/d)$. □

Corollary 12. *If* $ca \equiv cb(mod\ n)$ *and* $gcd(c,n) = 1$, *then* $a \equiv b(mod\ n)$.

Corollary 13. *If* $ca \equiv cb(mod\ p)$ *and* $p \nmid c$, *where* p *is a prime number, then* $a \equiv b(mod\ p)$.

Proof. The conditions $p \nmid c$ and p a prime imply that $gcd(c,p) = 1$. Then by Corollary 12, $a \equiv b(mod\ p)$. □

BINARY AND DECIMAL REPRESENTATIONS OF INTEGERS.

One of the more interesting applications of congruence theory involves finding special criteria under which a given integer is divisible by another integer. At their heart, these divisibility tests depend on the notational system used to assign "names" to integers and, more particularly, to the fact that 10 is taken as the base for our number system. Let us, therefore, start by showing that, given an integer $b > 1$, any positive integer N can be written uniquely in terms of powers of b as

$$N = a_m b^m + a_{m-l} b^{m-l} + \cdots + a_2 b^2 + a_1 b + a_0$$

where the coefficients a_k can take on the b different values $0, 1, 2, \cdots, b-1$. For the Division Algorithm yields integers q_1 and a_0 satisfying

$$N = q_1 b + a_0 \qquad 0 \leq a_0 < b.$$

If $q_1 \geq b$, we can divide once more, obtaining

$$q_1 = q_2 b + a_1 \qquad 0 \geq a_1 < b.$$

Now substitute for q_1 in the earlier equation to get

$$N = (q_2 b + a_1)b + a_0 = q_2 b^2 + a_1 b + a_0.$$

Because $N > q_1 > q_2 > \cdots \geq 0$ is a strictly decreasing sequence of integers, this process must eventually terminate, say, at the $(m-1)$th stage, where

$$q_{m-1} = q_m b + a_{m-1} \qquad 0 \leq a_{m-1} < b$$

and $0 \leq q_m < b$. Setting $a_m = q_m$, we reach the representation

$$N = a_m b^m + a_{m-l} b_{m-l} + \cdots + a_1 b + a_0.$$

which was our aim. To show uniqueness, refer text.

Theorem 18. *Let $P(x) = \sum_{k=0}^{m} c_k x^k$ be a polynomial function of x with integral coefficients c_k. If $a \equiv b(mod\ n)$, then $P(a) \equiv P(b)(mod\ n)$.*

Proof. Please refer text. □

Theorem 19. *If a is a solution of $P(x) \equiv 0(mod\ n)$ and $a \equiv b(mod\ n)$, then b also is a solution.*

Proof. From the last theorem, it is known that $P(a) \equiv P(b)(mod\ n)$. Hence, if a is a solution of $P(x) \equiv 0(mod\ n)$, then $P(b) \equiv P(a) \equiv 0(mod\ n)$, making b a solution. □

Theorem 20. *Let $N = a_m 10^m + a_{m-1}, 10^{m-1} + \cdots + a_1 10 + a_0$ be the decimal expansion of the positive integer N, $0 \leq a_k < 10$, and let $S = a_0 + a_1 + \cdots + a_m$. Then $9|N$ if and only if $9|S$.*

Proof. Please refer text. □

Theorem 21. *Let $N = a_m 10^m + a_{m-1} 10^{m-1} + \cdots + a_1 10 + a_0$ be the decimal expansion of the positive integer N, $0 \leq a_k < 10$, and let $T = a_0 - a_1 + a_2 - \cdots + (-1)^m a_m$. Then $11|N$ if and only if $11|T$.*

Proof. Please refer text. □

LINEAR CONGRUENCES AND THE CHINESE REMAINDER THEOREM.

Theorem 22. *The linear congruence $ax \equiv b(mod\ n)$ has a solution if and only if $d|b$, where $d = gcd(a, n)$. If $d|b$, then it has d mutually incongruent solutions modulo n.*

Theorem 23. *Chinese Remainder Theorem: Let $n_1, n_2, \cdots, n$, be positive integers such that $gcd(n_i, n_j) = 1$ for $i \neq j$. Then the system of linear congruences*

$$\begin{aligned} x &\equiv a_1 (mod\ n_1) \\ x &\equiv a_2 (mod\ n_2) \\ &\vdots \\ x &\equiv a_r (mod\ n_r) \end{aligned}$$

has a simultaneous solution, which is unique modulo the integer $n_1, n_2, \cdots, n_r$.

Problem 1: Solve

$$\begin{aligned} x &\equiv 2 (mod\ 3) \\ x &\equiv 3 (mod\ 5) \\ x &\equiv 2 (mod\ 7). \end{aligned}$$

Solution: We have $n = 3\ .\ 5\ .7 = 105$ and

$$N_1 = \frac{n}{3} = 35 \qquad N_2 = \frac{n}{5} = 21 \qquad N_3 = \frac{n}{7} = 15$$

Now the linear congruences

$$35x \equiv 1 (mod\ 3) \qquad 21x \equiv 1 (mod\ 5) \qquad 15x \equiv 1 (mod\ 7)$$

are satisfied by $x_1 = 2$, $x_2 = 1$, $x_3 = 1$, respectively. Thus, a solution of the system is given by

$$x = 2352 + 3211 + 2151 = 233$$

Modulo 105, we get the unique solution $x = 233 = 23 (mod\ 105)$.

MODULE II

Chapter 4

FERMAT'S THEOREM

FERMAT'S LITTLE THEOREM AND PSEUDOPRIMES

The most significant of Fermat's correspondents in number theory was Bernhard Frenicle de Bessy (1605-1675), an official at the French mint who was renowned for his gift of manipulating large numbers. (Frenicle 's facility in numerical calculation is revealed by the following incident: On hearing that Fermat had proposed the problem of finding cubes that when increased by their proper divisors become squares, as is the case with $7^3+(1+7+72) = 20^2$, he immediately gave four different solutions, and supplied six more the next day.) Though in no way Fermat's equal as a mathematician, Frenicle alone among his contemporaries could challenge Fermat in number theory and Frenicle's challenges had the distinction of coaxing out of Fermat some of his carefully guarded secrets. One of the most striking is the theorem that states: If p is a prime and a is any integer not divisible by p, then p divides $a^{p-1} - 1$. Fermat communicated the result in a letter to Frenicle dated October 18, 1640, along with the comment, "I would send you the demonstration, if I did not fear its being too long." This theorem has since become known as "Fermat's Little Theorem," or just "Fermat's Theorem," to distinguish it from Fermat's "Great" or "Last Theorem," which is the subject of Chapter 12. Almost 100 years were to elapse before Euler published the first proof of the little theorem in 1736. Leibniz, however, seems not to have received his share of recognition, for he left an identical argument in an unpublished manuscript sometime before 1683.
We now proceed to a proof of Fermat's theorem.

Theorem 24. *Fermat's theorem. Let p be a prime and suppose that $p|a$. Then*

$$a^{P-1} \equiv 1 (mod\ p).$$

Proof. We begin by considering the first $p-1$ positive multiples of a; that is, the integers

$$a, 2a, 3a, \cdots, (p-1)a$$

None of these numbers is congruent modulo p to any other, nor is any congruent to zero. Indeed, if it happened that

$$ra \equiv sa(mod\ p) \qquad 1 \leq r < s \leq p-1$$

then a could be canceled to give $r \equiv s(mod\ p)$, which is impossible. Therefore, the previous set of integers must be congruent modulo p to $1, 2, 3, \cdots, p-1$, taken in some order. Multiplying all these congruences together, we find that

$$a.2a.3a \cdots (p-1)a \equiv 1.2.3 \cdots (p-1)(mod\ p)$$

whence

$$a^{p-1}(p-1)! \equiv (p-1)!(mod\ p).$$

Once $(p-1)!$ is canceled from both sides of the preceding congruence (this is possible because since $p|(p-1)!$), our line of reasoning culminates in the statement that $a^{p-l} = 1(mod\ p)$, which is Fermat's theorem. □

This result can be stated in a slightly more general way in which the requirement that $p \nmid a$ is dropped.

Corollary 14. *If p is a prime, then $a^p = a(mod\ p)$ for any integer a.*

Proof. When $p|a$, the statement obviously holds; for, in this setting, $a^p \equiv 0 \equiv a(mod\ p)$. If $p|a$, then according to Fermat's theorem, we have $a^p - l \equiv 1(mod\ p)$. When this congruence is multiplied by a, the conclusion $a^p \equiv a(mod\ p)$ follows. □

Lemma 2. *If p and q are distinct primes with $a^p \equiv a(mod\ q)$ and $a^q \equiv a(mod\ p)$, then $a^{pq} = a(mod\ pq)$.*

Proof. The last corollary tells us that $(a^q)^p = a^q(mod\ p)$, whereas $a^q = a(mod\ p)$ holds by hypothesis. Combining these congruences, we obtain $a^{pq} = a(mod\ p)$. or, in different terms, $p|a^{pq} - a$. In similar manner, $q|a^{pq} - a$. Therefore $pq|a^{pq} - a$, that is, $a^{pq} = a(mod\ pq)$. □

Definition 9. *A composite integer n is called pseudoprime whenever $n|2n-2$. It can be shown that there are infinitely many pseudoprimes, the smallest four being 341, 561, 645, and 1105.*

Theorem 25. *If n is an odd pseudoprime, then $M_n = 2^n - 1$ is a larger one.*

Proof. Please refer text. □

WILSON'S THEOREM

We now turn to another milestone in the development of number theory. In his Meditationes Algebraicae of 1770, the English mathematician Edward Waring (1734-1798) announced several new theorems. Foremost among these is an interesting property of primes reported to him by one of his former students, a certain John Wilson. The property is the following: If p is a prime number, then p divides $(p-1)!+1$. Wilson appears to have guessed this on the basis of numerical computations; at any rate, neither he nor Waring knew how to prove it. Confessing his inability to supply a demonstration, Waring added, "Theorems of this kind will be very hard to prove, because of the absence of a notation to express prime numbers." (Reading the passage, Gauss uttered his telling comment on "notationes versus notiones," implying that in questions of this nature it was the notion that really mattered, not the notation.) Despite Waring's pessimistic forecast, soon afterward Lagrange (1771) gave a proof of what in literature is called "Wilson's theorem" and observed that the converse also holds. Perhaps it would be more just to name the theorem after Leibniz, for there is evidence that he was aware of the result almost a century earlier, but published nothing on the subject.
Now we give a proof of Wilson's theorem.

Theorem 26. *Wilson. If p is a prime, then $(p-1)! \equiv -1(mod\ p)$.*

Proof. When $p = 2$ and $p = 3$ are trivial, let us take $p > 3$. Suppose that a is any one of the $p-1$ positive integers $1, 2, 3, \cdots, p-1$ and consider the linear congruence $ax \equiv 1(mod\ p)$. Then $gcd(a, p) = 1$. Therefore this congruence admits a unique solution modulo p; hence, there is a unique integer a', with $1 \leq a' \leq p-1$, satisfying $aa' \equiv 1(mod\ p)$. Because p is prime, $a = a'$ if and only if $a = 1$ or $a = p-1$. Indeed, the congruence $a^2 \equiv 1(mod\ p)$ is equivalent to $(a-1).(a+1) = 0(mod\ p)$. Therefore, either $a-1 \equiv 0(mod\ p)$, in which case $a = 1$, or $a+1 \equiv 0(mod\ p)$, in which case $a = p-1$.
If we omit the numbers 1 and $p-1$, the effect is to group the remaining integers $2, 3, \cdots, p-2$ into pairs a, a', where $a \neq a'$, such that their product $aa' \equiv 1(mod\ p)$. When these $(p-3)/2$ congruences are multiplied together and the factors rearranged, we get

$$2.3 \cdots (p-2) \equiv 1(mod\ p)$$

or rather

$$(p-2)! \equiv 1(mod\ p)$$

Now multiply by $p-1$ to obtain the congruence

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

this completes the proof. □

Theorem 27. *The quadratic congruence $x^2+1 \equiv 0(mod\ p)$, where p is an odd prime, has a solution if and only if $p \equiv 1(mod 4)$.*

Proof. Please refer text. □

Chapter 5

NUMBER-THEORETIC FUNCTIONS

THE SUM AND NUMBER OF DIVISORS

Certain functions are found to be of special importance in connection with the study of the divisors of an integer. Any function whose domain of definition is the set of positive integers is said to be a number-theoretic (or arithmetic)function. Although the value of a number-theoretic function is not required to be a positive integer or, for that matter, even an integer, most of the number-theoretic functions that we shall encounter are integer-valued. Among the easiest to handle, and the most natural, are the functions τ and σ.

Definition 10. *Given a positive integer n, let $\tau(n)$ denote the number of positive divisors of n and $\sigma(n)$ denote the sum of these divisors.*

For an example of these notions, consider $n = 12$. Because 12 has the positive divisors 1, 2, 3, 4, 6, 12, we find that $\tau(12) = 6$ and $\sigma(12) = 1 + 2 + 3 + 4 + 6 + 12 = 28$ For the first few integers,

$$\tau(l) = 1\tau(2) = 2\tau(3) = 2\tau(4) = 3\tau(5) = 2\tau(6) = 4, \cdots$$

and

$$\sigma(1) = 1, \sigma(2) = 3, \sigma(3) = 4, \sigma(4) = 7, \sigma(5) = 6, \sigma(6) = 12, \cdots$$

It is not difficult to see that $\tau(n) = 2$ if and only if n is a prime number; also, $\sigma(n) = n + 1$ if and only if n is a prime.

Theorem 28. *If $n = p_1^{k_1} p_2^{k_2} \cdots p_r^{k_r}$ is the prime factorization of $n > 1$, then the positive divisors of n are precisely those integers d of the form*

$$d = p_1^{a_1} p_2^{a_2} \cdots p_r^{a_r}$$

where $0 \leq a_i \leq k_i \quad (i = 1, 2, ..., r)$.

Proof. Please refer text. □

Theorem 29. *If $n = p_1^{k_1} p_2^{k_2} \cdots p_r^{k_r}$ is the prime factorization of $n > 1$, then*

1. *$\tau(n) = (k_1 + 1)(k_2 + 1) \cdots (k_r + 1)$, and*
2. *$\sigma(n) = \frac{p_1^{k_1+1}-1}{p_1-1} \ \frac{p_2^{k_2+1}-1}{p_2-1} \ \frac{p_r^{k_r+1}-1}{p_r-1}$*

Proof. Please refer text. □

Problem: Find the number of positive divisors and their sum of 180.
Solution: The number $180 = 2^2 \,.\, 3^2 \,.\, 5$ has

$$\tau(180) = (2 + 1)(2 + 1)(1 + 1) = 18$$

positive divisors. These are integers of the form

$$2^{a_1} \,.\, 3^{a_2} \,.\, 5^{a_3}$$

where $a_1 = 0, 1, 2$; $a_2 = 0, 1, 2$; and $a_3 = 0, 1$. Specifically, we obtain $1, 2, 3, 4, 5, 6, 9, 10, 12, 15, 18, 20, 30, 36, 45, 60, 90, 180$.
The sum of these integers is

$$\sigma(180) = \frac{2^3 - 1}{2 - 1} \ \frac{3^3 - 1}{3 - 1} \ \frac{5^2 - 1}{5 - 1} = \frac{7}{1} \ \frac{26}{2} \frac{24}{4} = 7 \,.\, 13 \,.\, 6 = 546$$

Definition 11. *A number-theoretic function f is said to be multiplicative if*

$$f(mn) = f(m)f(n)$$

whenever $gcd(m, n) = 1$.

Example: The functions τ and σ are both multiplicative functions.

THE GREATEST INTEGER FUNCTION

The greatest integer or "bracket" function [ ] is especially suitable for treating divisibility problems. Although not strictly a number-theoretic function, its study has a natural place in this chapter.

Definition 12. *For an arbitrary real number x, we denote by $[x]$ the largest integer less than or equal to x; that is, $[x]$ is the unique integer satisfying $x - 1 < [x] \leq x$.*

$$[-3/2] = -2 \qquad [\sqrt{2}] = 1 \qquad [1/3] = 0 \qquad [\pi] = 3[-\pi] = -4$$

The important observation to be made here is that the equality $[x] = x$ holds if and only if x is an integer. Also from the definition we have any real number x can be written as

$$x = [x] + \theta$$

for a suitable choice of θ with $0 \leq \theta < 1$.

Results:

1. If n is a positive integer and p a prime, then the exponent of the highest power of p that divides $n!$ is

$$\sum_{k=1}^{\infty} \left[\frac{n}{p^k} \right]$$

   where the series is finite, because $[n/p^k] = 0$ for $p^k > n$.

2. If n and r are positive integers with $1 \leq r < n$, then the binomial coefficient

$$\left(\begin{array}{c} n \\ r \end{array} \right) = \frac{n!}{r!(n-r)!}$$

   is also an integer.

3. For a positive integer r, the product of any r consecutive positive integers is divisible by $r!$.

4. Let f and F be number-theoretic functions such that

$$F(n) = \sum_{d|n} f(d).$$

   Then, for any positive integer N,

$$\sum_{n=1}^{N} F(n) = \sum_{k=1}^{N} f(k) \left[\frac{N}{k} \right].$$

5. If N is a positive integer, then

$$\sum_{n=1}^{N} \tau(n) = \sum_{n=1}^{N} \left[\frac{N}{n} \right].$$

6. If N is a positive integer, then

$$\sum_{n=1}^{N} \sigma(n) = \sum_{n=1}^{N} n \left[\frac{N}{n} \right].$$

Chapter 6

EULER'S GENERALIZATION OF FERMAT'S THEOREM

EULER'S Pill-FUNCTION

This chapter deals with that part of the theory arising out of the result known as Euler's Generalization of Fermat's Theorem. In a nutshell, Euler extended Fermat's theorem, which concerns congruences with prime moduli, to arbitrary moduli. While doing so, he introduced an important number-theoretic function:

Definition 13. *For $n \geq 1$, let $\phi(n)$ denote the number of positive integers not exceeding n that are relatively prime to n.*

As an illustration of the definition, we find that $\phi(30) = 8$; for, among the positive integers that do not exceed 30, there are eight that are relatively prime to 30; specifically,

$$1, 7, 11, 13, 17, 19, 23, 29.$$

Similarly, for the first few positive integers, $\phi(l) = 1, \phi(2) = 1, \phi(3) = 2, \phi(4) = 2, \phi(5) = 4, \phi(6) = 2, \phi(7) = 6, \cdots$
Notice that $\phi(1) = 1$, because $gcd(1, 1) = 1$. In the event $n > 1$, then $gcd(n, n) = n \neq 1$, so that $\phi(n)$ can be characterized as the number of integers less than n and relatively prime to it. The function $\phi(n)$ is usually called the Euler phi-function (sometimes, the indicator or totient) after its originator; the functional notation $\phi(n)$, however, is credited to Gauss.
If n is a prime number, then every integer less than n is relatively prime to it; whence, $\phi(n) = n - 1$. On the other hand, if $n > 1$ is composite, then n has a divisor d such that $1 < d < n$. It follows that there are at least two

integers among $1, 2, 3, \cdots, n$ that are not relatively prime to n, namely, d and n itself. As a result, $\phi(n) \leq n - 2$. This proves that for $n > 1$,

$$\phi(n) = n - 1 \quad \textit{if and only if } n \textit{ is prime.}$$

Results:

1. If p is a prime and $k > 0$, then
$$\phi(p^k) = p^k - p^{k-1} = p^k \left(1 - \frac{1}{p}\right).$$
2. Given integers a, b, c, $gcd(a, bc) = 1$ if and only if $gcd(a, b) = 1$ and $gcd(a, c) = 1$.
3. The function ϕ is a multiplicative function.
4. If the integer $n > 1$ has the prime factorization $n = p_1^{k_1} p_2^{k_2} \cdots p_r^{k_r}$, then
$$\begin{aligned}\phi(n) &= (p_1^{k_1} - p_1^{k_1-1})(p_2^{k_2} - p_2^{k_2-1}) \cdots (p_r^{k_r} - p_r^{k_r-1}) \\ &= n \left(1 - \frac{1}{p_1}\right)\left(1 - \frac{1}{p_2}\right) \cdots \left(1 - \frac{1}{p_r}\right)\end{aligned}$$
5. For $n > 2$, $\phi(n)$ is an even integer.

Problem 1: Find $\phi(16)$.
Solution: $\phi(16) = \phi(2^4) = 2^4 - 2^3 = 16 - 8 = 8$.
Problem 2: Find $\phi(360)$.
Solution: The prime-power decomposition of 360 is $2^3 \cdot 3^2 \cdot 5$, therefore

$$\begin{aligned}\phi(360) &= 360(1 - \frac{1}{2})(1 - \frac{1}{3})(1 - \frac{1}{5}) \\ &= 360 \cdot \frac{1}{2} \cdot \frac{2}{3} \cdot \frac{4}{5} = 96.\end{aligned}$$

EULER'S THEOREM

Lemma 3. *Let $n > 1$ and $gcd(a, n) = 1$. If $a_1, a_2, \cdots, a_{\phi(n)}$ are the positive integers less than n and relatively prime to n, then*

$$aa_1,\ aa_2,\ \cdots a_{\phi(n)}$$

are congruent modulo n to $a_1, a_2, \cdots, a_{\phi(n)}$ in some order.

Proof. Observe that no two of the integers $a_1, a_2, \cdots, a_{\phi(n)}$ are congruent modulo n. For if $aa_i \equiv aa_j (mod\ n)$, with $1 \leq i < j \leq \phi(n)$, then the cancellation law yields $a_i \equiv a_j (mod\ n)$, and thus $a_i = a_j$, a contradiction. Furthermore, because $gcd(a_i, n) = 1$ for all i and $gcd(a, n) = 1$, this implies that each of the aa_i is relatively prime to n.
Fixing on a particular aa_i, there exists a unique integer b, where $0 \leq b < n$, for which $aa_i \equiv b (mod\ n)$. Because $gcd(b, n) = gcd(aa_i, n) = 1$, b must be one of the integers $a_1, a_2, \cdots, a_{\phi(n)}$. All told, this proves that the numbers $aa_1,\ aa_2,\ \cdots a_{\phi(n)}$ and the numbers $a_1, a_2, \cdots, a_{\phi(n)}$ are identical modulo n in a certain order. □

Theorem 30. *Euler. If $n \geq 1$ and $gcd(a, n) = 1$, then $a^{\phi}(n) = 1\ (mod\ n)$.*

Corollary 15. *Fermat. If p is a prime and $p \nmid a$, then $a^{p-1} \equiv 1 (mod\ p)$.*

Problem: Find the last two digits in the decimal representation of 3^{256}.
Solution: This is equivalent to obtaining the smallest nonnegative integer to which 3^{256} is congruent modulo 100. Because $gcd(3, 100) = 1$ and

$$\phi(100) = \phi(2^2 \,.\, 5^2) = 100 \left(1 - \frac{1}{2}\right)\left(1 - \frac{1}{5}\right) = 40.$$

Euler's theorem yields

$$3^{\phi(100)} \equiv 1 (mod\ 100)$$

$$3^4 0 \equiv 1 (mod\ 100)$$

By the Division Algorithm, $256 = 6 \,.\, 40 + 16$; whence

$$3^{256} \equiv 3^{6 \,\cdot\, 40+16} \equiv (3^{40})^6\ 3^{16} \equiv 3^{16} (mod\ 100)$$

and our problem reduces to one of evaluating 3^{16}, modulo 100. The method of successive squaring yields the congruences

$$3^2 \equiv 9 (mod\ 100) \qquad 3^4 \equiv 81 (mod\ 100)$$

$$3^8 \equiv 61 (mod\ 100) \qquad 3^{16} \equiv 21 (mod\ 100).$$

Hence the last two digits of 3^{256} is 21.

SOME PROPERTIES OF THE PHI-FUNCTION

The next theorem points out a curious feature of the phi-function; namely, that the sum of the values of $\phi(d)$, as d ranges over the positive divisors of n, is equal to n itself. This was first noticed by Gauss.

Theorem 31. *For each positive integer $n \geq 1$,*

$$n = \sum_{d|n} \phi(d)$$

the sum being extended over all positive divisors of n.

Theorem 32. *For $n > 1$, the sum of the positive integers less than n and relatively prime to n is $\frac{1}{2}n\phi(n)$.*

MODULE III

Chapter 7

RANK OF A MATRIX

Definition 14. *Let A be am $m \times n$ matrix. Then the matrix A or any matrix obtained by deleting some rows or columns of A is called sub-matrix of A.*

Definition 15. *Let A be am $m \times n$ matrix given by*

$$A = [a_{ij}]_{m\times n} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ & \cdot & \cdot & \cdot \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

If we retain any t rows and t columns of A and deleting $m-t$ rows and $n-t$ columns, we obtain a $t \times t$ square sub-matrix of A. The determinant of this square sub-matrix of order t is called a minor of A of order t.

Definition 16. *The number r is called the rank of the matrix A if it satisfies the following properties:*

1. *There is at least one non-zero minor of order r.*
2. *Every minor of order $r+1$ is zero.*

Remarks:

- If the rank is r, then every minor of order $r+2$, being the sum of multiples of minors of order $r+1$, is zero. In fact, every minor of order greater than r (if it exists) will be zero. Therefor we can define rank as follows:
  The **rank of a matrix** A is the largest order of a nonzero minor of the matrix A. That is, there exists a nonzero minor of order L implies rank $\geq l$ and all minors of order $l+1$ are zero implies $\leq l$.

- The rank of a matrix A is denoted by $\rho(A)$.
- The rank of every nonsingular matrix of order n is n.
- The rank of any nonzero matrix is greater than or equal to 1.
- The rank of a zero matrix is taken as 0.

Problem 1: Find the rank of the matrix $A = \begin{bmatrix} 1 & 3 \\ 3 & 9 \end{bmatrix}$

Solution: Since $|A| = 0$, $\rho(A) = 1$

Problem 2: Find the rank of the matrix $A = \begin{bmatrix} 1 & 2 \\ 2 & 3 \\ 1 & 2 \\ 1 & 2 \end{bmatrix}$

Solution: Since there is no minors of order 4 and 3, and hence $\rho(A) < 3$. Now A has a minor $\begin{vmatrix} 1 & 2 \\ 2 & 3 \end{vmatrix} = -1 \neq 0$, and since its order is 2, $\rho(A) = 2$.

Problem 3: Find the rank of the matrix $A = \begin{bmatrix} 1 & 3 & 4 & 1 \\ 2 & 6 & 8 & 2 \\ 1 & 2 & 5 & 9 \\ 1 & 2 & 5 & 9 \end{bmatrix}$

Solution: Left as exercise.

Problem 4: Prove that the rank of matrix every element of which is unity is 1.

Solution; Since all elements are 1, square matrix of every order will have determinant 0, except the square matrix [1] of order 1.

Problem 5: Show that no skew-symmetric matrix can be of rank 1.

Solution: Let A be a skew-symmetric matrix. If A is zero matrix, then $\rho(A) = 0 \neq 1$. If A is nonzero matrix, then there exists at least one minor of the form $\begin{vmatrix} 0 & a \\ -a & 0 \end{vmatrix} = a^2 \neq 0$. Hence rank of A is not 1.

Problem 6: The rank of the transpose of a matrix is same that of the original matrix.

Solution: Left as exercise.

Chapter 8

ELEMENTARY TRANSFORMATIONS AND EQUIVALENT MATRICES

Definition 17. *An elementary transformation is an operation of any one of the following types:*

1. *the interchange of any two rows (columns)*
2. *the multiplication of the elements of any row (column) by any nonzero number.*
3. *the addition to the elements of any row (column), the corresponding elements of any other row (column) multiplied by any number.*

An elementary transformation is is called row transformation or a column transformation according as it applies to rows or columns.

Exercise: Find the inverse row operations corresponding to the elementary row (column)operations that have been dened just above.

Definition 18. *Equivalent Matrices: Two matrices are said to be equivalent if one can be obtained from the other by a nite number of elementary transformations.*

Exercise: Prove that equivalent matrices have same rank.

Determination of rank using elementary transformations

Theorem 33. *Every $m \times n$ matrix of rank r can by a sequence of elementary transformations br reduced to any one of the form*

$$\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} I_r & 0 \end{bmatrix}, \quad \begin{bmatrix} I_r \\ 0 \end{bmatrix}, \quad \begin{bmatrix} I_r \end{bmatrix},$$

where I_r is the identity matrix of order r and each $0's$ are zeromatrices of appropriate orders.

Definition 19. *One of the forms*

$$\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} I_r & 0 \end{bmatrix}, \quad \begin{bmatrix} I_r \\ 0 \end{bmatrix}, \quad \begin{bmatrix} I_r \end{bmatrix},$$

to which a given matrix A of rank r can be reduced by elementary transformations is known as normal form of the matrix.

Theorem 34. *If an $m \times n$ matrix A can be reduced to any one of the form*

$$\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} I_r & 0 \end{bmatrix}, \quad \begin{bmatrix} I_r \\ 0 \end{bmatrix}, \quad \begin{bmatrix} I_r \end{bmatrix},$$

by a sequence of elementary transformations, then the rank of A is r.

Problem: 1 Reduce the matrix $A = \begin{bmatrix} 1 & 2 & -1 \\ 2 & 4 & -2 \\ 3 & 6 & -3 \end{bmatrix}$ to its normal form hence find rank.

Solution:

$$\begin{aligned} A &= \begin{bmatrix} 1 & 2 & -1 \\ 2 & 4 & -2 \\ 3 & 6 & -3 \end{bmatrix} \\ &\sim \begin{bmatrix} 1 & 2 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{array}{l} \text{by } R_2 \to R_2 - 2R_1 \\ \text{by } R_3 \to R_3 - 3R_1 \end{array} \\ &\sim \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{array}{l} \text{by } C_2 \to C_2 - 2C_1 \\ \text{by } C_3 \to C_3 + C_1 \end{array} \\ &\sim \begin{bmatrix} I_1 & 0 \\ 0 & 0 \end{bmatrix} . \end{aligned}$$

Hence rank of A is 1.

Problem: 2 Reduce normal form the matrix

$$A = \begin{bmatrix} 0 & 3 & 1 & 3 & 0 \\ 0 & 0 & 1 & 0 & 5 \\ 2 & 0 & 1 & 0 & 5 \\ 7 & 1 & 1 & 1 & 7 \end{bmatrix}.$$

Elementary Matrices

Definition 20. *A matrix obtained from a identity matrix by subjecting it to any of the elementary transformations is called an elementary matrix.*

Results

1. Every elementary row transformation of a product of two matrices can be effected by subjecting the pre-factor to the same row transformation.
2. Every elementary column transformation of a product of two matrices can be effected by subjecting the post-factor to the same column transformation.
3. Every elementary row transformation of a matrix can be brought about by pre-multiplication with the corresponding elementary matrix.
4. Every elementary column transformation of a matrix can be brought about by post-multiplication with the corresponding elementary matrix.
5. The inverse of an elementary matrix is also an elementary matrix of the same type.
6. Every elementary matrix is nonsingular.
7. Corresponding to a matrix A of rank r there exist nonsingular matrices P and Q such that
$$PAQ = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$$
8. Every nonsingular matrix is a product of elementary matrices.
9. The rank of matrix is not altered by pre-multiplication or post-multiplication with any nonsingular matrix.
10. The rank of the product of two matrices cannot exceed the rank of either matrix.

Inverse using Elementary Transformations

We have every nonsingular matrix can by a series of elementary transformations, be reduced to the unit(identity) matrix. This gives us a method for computing the inverse of any nonsingular matrix A. The procedure is as follows:

Step 1 Write

$$A = IA$$

Step 2 go on performing suitable elementary row transformations on the matrix A on the left and the pre-factor I of A on the right till we reach the result

$$I = BA.$$

Step 3 Then B is the inverse of A.

Problem 1: Compute the inverse of

$$A = \begin{bmatrix} 1 & 6 & 4 \\ 0 & 2 & 3 \\ 0 & 1 & 2 \end{bmatrix}.$$

Solution: We write

$$A = IA$$

That is,

$$\begin{bmatrix} 1 & 6 & 4 \\ 0 & 2 & 3 \\ 0 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} A.$$

Performing $R_2 \to (\frac{1}{2})R_2$ on the matrix A on the left and the pre-factor I of A on the right, we obtain

$$\begin{bmatrix} 1 & 6 & 4 \\ 0 & 1 & \frac{3}{2} \\ 0 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} A.$$

Performing $R_1 \to R_1 + (-6)R_2$, $R_3 \to R_3 + (-1)R_2$ on the matrix A on the left and the pre-factor I of A on the right, we obtain

$$\begin{bmatrix} 1 & 0 & -5 \\ 0 & 1 & \frac{3}{2} \\ 0 & 0 & \frac{1}{2} \end{bmatrix} = \begin{bmatrix} 1 & -3 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & -\frac{1}{2} & 1 \end{bmatrix} A.$$

Performing $R_3 \to 2R_3$ on the matrix A on the left and the pre-factor I of A on the right, we obtain

$$\begin{bmatrix} 1 & 0 & -5 \\ 0 & 1 & \frac{3}{2} \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -3 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & -1 & 2 \end{bmatrix} A.$$

Performing $R_1 \to R_1 + 5R_3$, $R_2 \to R_2 + (-\frac{3}{2})R_3$ on the matrix A on the left and the pre-factor I of A on the right, we obtain

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -8 & 10 \\ 0 & 2 & -3 \\ 0 & -1 & 2 \end{bmatrix} A.$$

Hence

$$A^{-1} = \begin{bmatrix} 1 & -8 & 10 \\ 0 & 2 & -3 \\ 0 & -1 & 2 \end{bmatrix}.$$

Problem 2: Compute the inverse of

$$A = \begin{bmatrix} -1 & -3 & 3 & -1 \\ 1 & 1 & -1 & 0 \\ 2 & -5 & 2 & -3 \\ -5 & 1 & 0 & 1 \end{bmatrix}.$$

Chapter 9

SYSTEM OF LINEAR EQUATIONS

A linear equation in variables $x_1, x_2, \cdots, x_n$ is an equation of the form

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b,$$

where $a_1, a_2, \cdots a_n$ and b are constant real or complex numbers. The constant a_i is called the coefficient of x_i and b is called the constant term of the equation.
A system of linear equations (or linear system) is a finite collection of linear equations in same variables. For instance, a linear system of m equations in n variables $x_1, x_2, \cdots, x_n$ can be written as

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n} = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n} = b_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn} = b_m \end{cases} \tag{9.1}$$

A solution of a linear system is a n-tuple $(s_1, s_2, \cdots, s_n)$ of numbers that makes each equation a true statement when the values $s_1, s_2, \cdots, s_n$ are substituted for $x_1, x_2, \cdots, x_n$, respectively. The set of all solutions of a linear system is called the solution set of the system.
Any system of linear equations has one of the following exclusive conclusions.
(a) No solution.
(b) Unique solution.
(c) Infinitely many solutions.
A linear system is said to be consistent if it has at least one solution and is said to be inconsistent if it has no solution.
The system of equations (9.1) is said to be homogeneous if all b_j are zero;

otherwise, it is said to be non-homogeneous.
The system of equations (9.1) can be expressed as the single matrix equation

$$AX = B, \tag{9.2}$$

where

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ . & . & \cdots & . \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \qquad X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \qquad \text{and} \qquad B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}.$$

Any vector (column matrix) X that satisfies the matrix equation (9.2) is also the solution of the system.

Definition 21. *The matrix $[AB]$ which is obtained by placing the constant column matrix B to the right of the matrix A is called the augmented matrix. Thus the augmented matrix of the system $AX = B$ is*

$$[AB] = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ . & . & \cdots & . & . \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{bmatrix}$$

Theorem 35. *The system $AX = B$ is consistent if and only if A and $[AB]$ have the same rank.*

System of non-homogeneous Equations
If we are given with a system of m equations in n unknowns, proceed as follows:

1. Write down the corresponding matrix equation $AX = B$.
2. By elementary row transformations obtain row echelon matrix of the augmented matrix $[AB]$.
3. Examine whether the rank of A and the rank of $[AB]$ are the same or not.

Case 1 If rank of $A \neq$ rank of $[AB]$, then the system is inconsistent and has no solution.

Case 2 If rank of A = rank of $[AB]$, then the system is consistent.

Case 2a If rank of A = rank of $[AB] = n$ = number of unknowns, then the system has unique solution.

Case 2b If rank of A = rank of $[AB] < n$ = number of unknowns, then the system has infinitely many solutions. We assign arbitrary values to $(n - r)$ unknowns and determine the remaining r unknowns uniquely.

Problem 1: Test the following system of equations for consistence and solve it, if it is consistent.

$$\begin{array}{rcrcrcl} x & + & y & + & z & = & 6 \\ x & - & y & + & 2z & = & 5 \\ 3x & + & y & + & z & = & 8 \end{array} .$$

Solution: The matrix equation corresponding to the given system of equation is

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 2 \\ 3 & 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 6 \\ 5 \\ 8 \end{bmatrix}$$

The augmented matrix $[AB]$ is given by

$$[AB] = \begin{bmatrix} 1 & 1 & 1 & 6 \\ 1 & -1 & 2 & 5 \\ 3 & 1 & 1 & 8 \end{bmatrix}$$

By elementary row transformations, the above augmented matrix is equivalent to the matrix

$$\begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 3 \end{bmatrix} .$$

This implies that rank of $A = 3$ =rank of $[AB]$ = number of unknowns. Therefore the system has unique solution. The solution to the system of equations is $x = 1$; $y = 2$; $z = 3$.

Problem 2: Show that the system of equations

$$\begin{array}{rcrcrcl} x & + & 2y & + & z & = & 2 \\ 3x & + & y & - & 2z & = & 1 \\ 4x & - & 3y & - & z & = & 3 \\ 2x & + & 4y & + & 2z & = & 4 \end{array}$$

is consistent and hence solve.

Problem 3: Investigate for what values of a, b the system of equations

$$\begin{array}{ccccccc} x & + & y & + & 2z & = & 2 \\ 2x & - & y & + & 3z & = & 10 \\ 5x & - & y & + & az & = & b \end{array}$$

have

1. no solution;
2. unique solution; and
3. an infinite number of solutions.

Solution: Here the augmented matrix is

$$[AB] = \begin{bmatrix} 1 & 1 & 2 & 2 \\ 2 & -1 & 3 & 10 \\ 5 & -1 & a & b \end{bmatrix}$$

By elementary row transformations, the above augmented matrix is equivalent to the matrix

$$[AB] \sim \begin{bmatrix} 1 & 0 & \frac{5}{3} & 4 \\ 0 & 1 & \frac{1}{3} & -2 \\ 0 & 0 & a-8 & b-22 \end{bmatrix}$$

1. the given system of equations have no solution if and only if rank of $A \neq$ rank of the augmented matrix $[AB]$. From the matrix equivalent to $[AB]$, it can be seen that this is possible if and only if

   $$a - 8 = 0 \text{ and } b - 22 \neq 0.$$

   That is, if and only if

   $$a = 8 \text{ and } b \neq 22.$$

   That is, the system has no solution if $a = 8$ and $b \neq 22$.

2. the given system of equations have unique solution if and only if rank of $A =$ rank of the augmented matrix $[AB] = 3 =$ number of unknowns. From the matrix equivalent to $[AB]$, it can be seen that this is possible if and only if

   $$a - 8 \neq 0 \text{ and } b - 22\text{has any value.}$$

   That is, if and only if

   $$a \neq 8 \text{ and } b\text{take any value.}$$

   That is, the system has unique solution if $a \neq 8$ and b take any value.

3. the given system of equations have infinitely many solutions if and only if rank of A = rank of the augmented matrix $[AB] < 3$ = number of unknowns. From the matrix equivalent to $[AB]$, it can be seen that this is possible if and only if

$$a - 8 = 0 \text{ and } b - 22 = 0.$$

That is, if and only if

$$a = 8 \text{ and } b = 22.$$

That is, the system has infinitely many solutions if $a = 8$ and $b = 22$.

Problem 4: Find the values λ and μ in order that the system of equations

$$\begin{array}{ccccccc} x & + & y & + & z & = & 6 \\ x & + & 2y & + & 3z & = & 10 \\ x & + & 2y & + & \lambda z & = & \mu \end{array}$$

have

1. no solution;
2. unique solution; and
3. an infinite number of solutions.

Note: The system of homogeneous is always consistent, because $X = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$ is obviously a solution, called trivial solution.

Definition 22. *The subspace generated by the vectors X such that*

$$AX = O$$

called the column null space of the $m \times n$ matrix A and its dimension $n - r$ is called the column nullity of the matrix.

In view of the definition, we have

$$\text{rank+column nullity= number of columns.}$$

Definition 23. *The subspace generated by the vectors X such that*

$$YX = O$$

called the row null space of the $m \times n$ matrix A and its dimension $n - r$ is called the row nullity of the matrix.

In view of the definition, we have

rank+row nullity= number of rows.

Theorem 36. *The row nullity and the column nullity of the square matrix are equal.*

MODULE IV

Chapter 10

CHARACTERISTIC ROOTS AND CHARACTERISTIC VECTORS OF A MATRIX

Definition 24. *A nonzero vector X is an eigenvector (or characteristic vector) of a square matrix A if there exists a scalar λ such that $AX = \lambda X$. Then λ is an eigenvalue (or characteristic value) of A.*

Let X be an eigenvector of the matrix A. Then there must exist an eigenvalue λ such that $AX = \lambda X$ or, equivalently, $AX - \lambda X = 0$ or $(A - \lambda I)X = 0$. If we define a new matrix $B = A - \lambda I$, then $BX = 0$. If B has an inverse then $X = B^{-1}0 = 0$. But an eigenvector cannot be zero. Thus, it follows that X will be an eigenvector of A if and only if B does not have an inverse, or equivalently $\det(B) = 0$, or $\det(A - \lambda I) = 0$. This is called the characteristic equation of A. Its roots determine the eigenvalues of A.
Problem 1: Find the eigenvalues of

$$A = \begin{bmatrix} 1 & 1 & 2 \\ 0 & 2 & 2 \\ -1 & 1 & 3 \end{bmatrix}.$$

Solution: The characteristic equation is $|A - \lambda I| = 0$. That is,

$$\begin{vmatrix} 1-\lambda & 1 & 2 \\ 0 & 2-\lambda & 2 \\ -1 & 1 & 3-\lambda \end{vmatrix} = 0,$$

which on simplification yields,

$$\lambda^3 - 6\lambda^2 + 11\lambda - 6 = 0.$$

By inspection $\lambda = 1$ is an eigenvalue. Divide $\lambda^3 - 6\lambda^2 + 11\lambda - 6$ by $(\lambda - 1)$, we get

$$(\lambda - 1)(\lambda^2 - 5\lambda + 6) = 0.$$

That is,

$$(\lambda - 1)(\lambda - 3)(\lambda - 2) = 0.$$

Therefore the eigenvalues are $\lambda = 1,\ \lambda = 2,\ \lambda = 3$. **Results:**

1. The eigenvalues of a diagonal matrix are the same as its diagonal elements.

2. The eigenvalues of a triangular matrix are the same as its diagonal elements.

3. A matrix A and its transpose A^T have the same characteristic roots.

4. If A and P be square matrices of the same order and if P be invertible, then the matrices A and $P^{-1}AP$ have the same characteristic roots.

5. Two similar matrices have the same characteristic roots.

Problem 1: Prove that 0 is a characteristic root of a matrix if and only if the matrix is singular.
Solution: The characteristic equation is $|A - \lambda I| = 0$. Suppose $\lambda = 0$ is a characteristic root of A, then we have $|A| = 0$. That is, A is singular. Conversely suppose the matrix A is singular, then $|A| = 0$. This implies that $|A - 0I| = 0$. That is, 0 is a characteristic root of A.
Problem 2: If λ is a characteristic root of a nonsingular matrix, then prove that $\lambda \neq 0$.
Problem 3: If λ is a characteristic root of a nonsingular matrix A, then λ^{-1} is a characteristic root of A^{-1}.
Solution: If A is nonsingular $\lambda \neq 0$.

$$\begin{array}{rcl} A - \lambda I & = & A - \lambda I \\ = & = & A - \lambda A A^{-1} \\ = & = & A(I - \lambda A^{-1}) \\ = & = & A(\lambda\lambda^{-1} I - \lambda A^{-1}) \\ = & = & A\lambda(\lambda^{-1} I - A^{-1}) \\ = & = & -\lambda A(A^{-1} - \lambda^{-1} I) \end{array}.$$

This implies that,

$$0 = |A - \lambda I| = |-\lambda A(A^{-1} - \lambda^{-1} I)|.$$

That is,

$$|(A^{-1} - \lambda^{-1}I)| = 0,$$

that is λ^{-1} is a characteristic root of A^{-1}.

Problem 4: Find the eigenvalues of the matrix $A = \begin{bmatrix} 1 & -2 \\ -5 & 4 \end{bmatrix}$ and hence find the matrix whose eigenvalues are $\frac{1}{6}$ and -1.

Problem 5: if the characteristic roots of A are $\lambda_1, \lambda_2, \cdots, \lambda_n$ then the characteristic roots of A^2 are $\lambda_1^2, \lambda_2^2, \cdots, \lambda_n^2$.

Solution: Suppose λ characteristic root of A then

$$|A - \lambda I| = 0.$$

Multiplied both side by $|A + \lambda I|$ we get,

$$|A - \lambda I||A + \lambda I| = 0|A + \lambda I| = 0.$$

That is,

$$|(A - \lambda I)(A + \lambda I)| = 0.$$

That is,

$$|(A^2 - \lambda^2 I)| = 0.$$

This implies that λ^2 is the characteristic roots of A^2.

Problem 6: Find the eigenvectors of the matrix $A = \begin{bmatrix} 10 & 3 \\ 4 & 6 \end{bmatrix}$.

Solution: The characteristic equation of A is $|A - \lambda I| = 0$. That is,

$$A = \begin{vmatrix} 10 - \lambda & 3 \\ 4 & 6 - \lambda \end{vmatrix} = 0.$$

On simplification we get,

$$\lambda^2 - 16\lambda + 48 = 0,$$

which gives the eigenvalues $\lambda = 4;\ \lambda = 12$.

The eigenvector X corresponding to $\lambda = 4$ is obtained by solving

$$(A - 4I)X = 0.$$

That is, by solving

$$\begin{bmatrix} 10 - 4 & 3 \\ 4 & 6 - 4 \end{bmatrix} X = 0.$$

That is, by solving

$$\begin{bmatrix} 6 & 3 \\ 4 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Applying row transformations, we finally get

$$\begin{bmatrix} 1 & \frac{1}{2} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Since the rank is 1 less than 2, the number of unknowns, we can assign arbitrary values to $n - r = 2 - 1 = 1$ unknowns, say $x_1 = a$, then $x_2 = -2a$. Similarly we can find The eigenvector X corresponding to $\lambda = 12$.

Chapter 11

CAYLEY-HAMILTON THEOREM

The Cayley-Hamilton Theorem states that every square matrix satisfies its own characteristic equation.

Problem 1: Find the characteristic equation of the matrix $A = \begin{bmatrix} 1 & 4 \\ 2 & 3 \end{bmatrix}$ and the verify Cayley-Hamilton Theorem.
Solution: The characteristic equation of A is $|A - \lambda I| = 0$. That is

$$A = \begin{vmatrix} 1-\lambda & 4 \\ 2 & 3-\lambda \end{vmatrix} = 0$$

On simplification we get,

$$\lambda^2 - 4\lambda - 5 = 0.$$

To verify Cayley-Hamilton Theorem for the matrix A, we have to verify that A satisfies the characteristic equation $\lambda^2 - 4\lambda - 5 = 0$. That is, we have to show that $A^2 - 4A - 5I = 0$.
Now $A^2 = \begin{bmatrix} 9 & 16 \\ 8 & 17 \end{bmatrix}$. Hence

$$A^2 - 4A - 5I = \begin{bmatrix} 9 & 16 \\ 8 & 17 \end{bmatrix} - 4\begin{bmatrix} 1 & 4 \\ 2 & 3 \end{bmatrix} - 5\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = 0.$$

Hence Cayley-Hamilton Theorem is verified for A.
Find the inverse of a matrix using the Cayley-Hamilton Theorem
Consider a square nonsingular matrix A of order n. Suppose the characteristic equation $|A - \lambda I| = 0$ of A is given by

$$a_0 + a_1\lambda + a_2\lambda^2 + \cdots + a_n\lambda^n = 0. \tag{11.1}$$

Now by Cayley-Hamilton Theorem, A satisfies equation (11.1), so that

$$a_0 I + a_1 A + a_2 A^2 + \cdots + a_n A^n = 0. \tag{11.2}$$

Pre-multiplying equation (11.2) by A^{-1}, we have

$$a_0 A^{-1} + a_1 I + a_2 A + \cdots + a_n A^{n-1} = 0, \tag{11.3}$$

which gives

$$A^{-1} = -\frac{1}{a_0}(a_1 I + a_2 A + \cdots + a_n A^{n-1}).$$

Problem 1: Using Cayley-Hamilton Theorem find the inverse of $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$.

Solution: The characteristic equation of A is $|A - \lambda I| = 0$. That is

$$A = \begin{vmatrix} 1-\lambda & 2 \\ 3 & 4-\lambda \end{vmatrix} = 0$$

On simplification we get,

$$\lambda^2 - 5\lambda - 2 = 0.$$

We have Cayley-Hamilton Theorem states that every square matrix satisfies its characteristic equation, hence

$$A^2 - 5A - 2I = 0.$$

Pre-multiplying the above equation by A^{-1}, we have

$$A - 5I - 2A^{-1} = 0.$$

That is,

$$A^{-1} = \frac{1}{2}(A - 5I) = \begin{bmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{bmatrix}.$$