Introduction to Number Theory

George Voutsadakis[1]

[1]Mathematics and Computer Science
Lake Superior State University

LSSU Math 400

Subsection 1

Foundations

Natural Numbers

- The set $1, 2, 3, \ldots$ of all natural numbers will be denoted by $\mathbb{N}$.

  $\mathbb{N}$ is a given set for which the Peano axioms are satisfied.
- They imply the following properties:
  - Addition and multiplication can be defined on $\mathbb{N}$, such that the commutative, associative and distributive laws are valid.
  - An ordering on $\mathbb{N}$ can be introduced so that either $m < n$ or $n < m$, for any distinct elements m, n in $\mathbb{N}$.
  - The principle of mathematical induction holds.
  - Every non-empty subset of $\mathbb{N}$ has a least element.

Integers, Rationals, Real and Complex Numbers

- We denote by $\mathbb{Z}$ the set of integers $0, \pm 1, \pm 2, \ldots$.
- We denote by $\mathbb{Q}$ the set of rationals, that is, the numbers $\frac{p}{q}$, with p in $\mathbb{Z}$ and q in $\mathbb{N}$.
- The construction, commencing with $\mathbb{N}$, of $\mathbb{Z}$, $\mathbb{Q}$ and then,
    - through Cauchy sequences, of the real numbers $\mathbb{R}$; and
    - through ordered pairs, of the complex numbers $\mathbb{C}$,

  forms the basis of mathematical analysis and it is assumed known.

Subsection 2

Division Algorithm

Divisibility

- Suppose that a, b are elements of $\mathbb{N}$.

  One says that b **divides** a (written $b \mid a$) if there exists an element c of $\mathbb{N}$, such that $a = bc$.
- In this case b is referred to as a **divisor** of a, and a is called a **multiple** of b.
- The relation $b \mid a$ is reflexive and transitive but not symmetric:

  In fact, if $b \mid a$ and $a \mid b$, then $a = b$.
- If $b \mid a$, then $b \leq a$; so a natural number has only finitely many divisors.
- The concept of divisibility is readily extended to $\mathbb{Z}$:

  If a, b are elements of $\mathbb{Z}$, with $b \neq 0$, then b is said to **divide** a if there exists c in $\mathbb{Z}$, such that $a = bc$.

The Division Algorithm

The Division Algorithm

For any a, b in $\mathbb{Z}$, with $b > 0$, there exist q, r in $\mathbb{Z}$, such that

$$a = bq + r \quad \text{and} \quad 0 \leq r < b.$$

- Suppose bq is the largest multiple of b that does not exceed a.

  Then the integer $r = a - bq$ is certainly non-negative.

  Since $b(q+1) > a$, we have $r < b$.
- The result remains valid for any integer $b \neq 0$ provided that the bound $r < b$ is replaced by $r < |b|$.

Subsection 3

Greatest Common Divisor

Greatest Common Divisor

- By the **greatest common divisor** of natural numbers a, b we mean an element d of $\mathbb{N}$, such that

$$d \mid a \quad \text{and} \quad d \mid b,$$

and, for every d' of $\mathbb{N}$,

$$d' \mid a \quad \text{and} \quad d' \mid b \quad \text{imply} \quad d' \mid d.$$

Existence of Greatest Common Divisors

Existence of Greatest Common Divisors

Given natural numbers a, b, the greatest common divisor d of a and b exists and is unique.

- Consider the set of all natural numbers of the form $ax + by$ with x, y in $\mathbb{Z}$. The set is not empty since, for instance, it contains a and b. Hence, there is a least member d, say. Now $d = ax + by$, for some integers x, y.
  - Clearly, every common divisor of a and b divides d.
  - By the division algorithm, $a = dq + r$, for some q, r in $\mathbb{Z}$, with $0 \leq r < d$. But, then, $r = a - dq = a - (ax + by)q = a(1 - qx) + b(-qy)$. From the minimal property of d, it follows that $r = 0$. So $d \mid a$.
  - Similarly, $d \mid b$

d is unique: Any other such number d' would divide d. Since, similarly, $d \mid d'$, we have $d = d'$.

Relatively Prime Pairs of Numbers

- We signify the greatest common divisor of a, b by (a, b).
- For any n in $\mathbb{N}$, the equation $ax + by = n$ is soluble in integers x, y if and only if (a, b) divides n.
- In the case $(a, b) = 1$, we say that a and b are **relatively prime** or **coprime** (or that a is **prime to** b).

  Then the equation $ax + by = n$ is always soluble.

Relatively Prime Numbers

- The concept of coprimality can be extended to more than two numbers.
- Any elements $a_1,\ldots,a_m$ of $\mathbb{N}$ have a greatest common divisor $d=(a_1,\ldots,a_m)$, such that

$$d=a_1x_1+\cdots+a_mx_m,$$

  for some integers $x_1,\ldots,x_m$.
- If $d=1$, we say that $a_1,\ldots,a_m$ are **relatively prime**.
- In the case of relatively prime $a_1,\ldots,a_m$, the equation

$$a_1x_1+\cdots+a_mx_m=n$$

  is always soluble.

Subsection 4

Euclid's Algorithm

Euclid's Algorithm

- Euclid's algorithm is a method for finding the greatest common divisor d of a, b:
  - By the division algorithm, there exist integers q_1, r_1, such that $a = bq_1 + r_1$ and $0 \leq r_1 < b$.
  - If $r_1 \neq 0$, then there exist integers q_2, r_2, such that $b = r_1 q_2 + r_2$, and $0 \leq r_2 < r_1$.
  - If $r_2 \neq 0$, then there exist integers q_3, r_3, such that $r_1 = r_2 q_3 + r_3$ and $0 \leq r_3 < r_2$.
  - Continuing thus, one obtains a decreasing sequence $r_1, r_2, \ldots$ satisfying $r_{j-2} = r_{j-1} q_j + r_j$.
  - The sequence terminates when $r_{k+1} = 0$, for some k, that is, when $r_{k-1} = r_k q_{k+1}$.

Euclid's Algorithm (Cont'd)

- Claim: $d = r_k$.

  Consider the equations

$$\begin{array}{ll} a = bq_1 + r_1, & 0 < r_1 < b; \\ b = r_1 q_2 + r_2, & 0 < r_2 < r_1; \\ r_1 = r_2 q_3 + r_3, & 0 < r_3 < r_2; \\ \cdots & \\ r_{k-2} = r_{k-1} q_k + r_k, & 0 < r_k < r_{k-1}; \\ r_{k-1} = r_k q_{k+1}. & \end{array}$$

  Every common divisor of a and b divides $r_1, r_2, \ldots, r_k$.

  Viewing the equations in the reverse order, r_k divides each r_j.

  Hence, r_k divides also b and a.

Applying Euclid's Algorithm I

- Euclid's algorithm enables the integers x, y, such that $d = ax + by$ to be explicitly calculated.

  Example: Take $a = 187$ and $b = 35$.

  Then, following Euclid, we have

  $$187 = 35 \cdot 5 + 12, \quad 35 = 12 \cdot 2 + 11, \quad 12 = 11 \cdot 1 + 1.$$

  Thus, we see that $(187, 35) = 1$. Moreover

  $$\begin{array}{rcl} 1 & = & 12 - 11 \cdot 1 = 12 - (35 - 12 \cdot 2) = 12 \cdot 3 - 35 \\ & = & (187 - 35 \cdot 5) \cdot 3 - 35 = 185 \cdot 3 + 35 \cdot (-16). \end{array}$$

  Hence, a solution of the equation $187x + 35y = 1$ in integers x, y is given by $x = 3, y = -16$.

Applying Euclid's Algorithm II

- Example: Take $a = 1000$ and $b = 45$. Then we get

$$1000 = 45 \cdot 22 + 10, \quad 45 = 10 \cdot 4 + 5, \quad 10 = 5 \cdot 2.$$

So $d = 5$.

The solutions to $ax + by = d$ can then be calculated from

$$\begin{array}{rcl} 5 & = & 45 - 10 \cdot 4 = 45 - (1000 - 45 \cdot 22)4 \\ & = & 1000 \cdot (-4) + 45 \cdot 89. \end{array}$$

This gives $x = -4$, $y = 89$.

Subsection 5

Fundamental Theorem

Prime Numbers and Prime Decomposition

- A natural number, other than 1, is called a **prime** if it is divisible only by itself and 1.

  The smallest primes are therefore given by $2, 3, 5, 7, 11, \ldots$
- Let n be any natural number other than 1.

  The least divisor of n that exceeds 1 is plainly a prime, say p_1.

  If $n \neq p_1$, then, similarly, there is a prime p_2 dividing $\frac{n}{p_1}$.

  If $n \neq p_1 p_2$, then there is a prime p_3 dividing $\frac{n}{p_1 p_2}$ and so on.

  After a finite number of steps, we obtain $n = p_1 \cdots p_m$;
- By grouping together we get the **standard factorization** (or **canonical decomposition**) $n = p_1^{j_1} \cdots p_k^{j_k}$, where $p_1, \ldots, p_k$ denote distinct primes and $j_1, \ldots, j_k$ are elements of $\mathbb{N}$.

Uniqueness of the Factorization

Uniqueness of Prime Factorization

The standard factorization is unique except for the order of the factors.

- If a prime p divides a product mn of natural numbers, then either p divides m or p divides n. If p does not divide m, then $(p, m) = 1$. Hence, there exist integers x, y, such that $px + my = 1$. Thus, we have $pnx + mny = n$. Hence, p divides n.

  More generally we conclude that if p divides $n_1 n_2 \cdots n_k$, then p divides n_ℓ, for some ℓ.

  Now suppose that, apart from the factorization $n = p_1^{j_1} \cdots p_k^{j_k}$, there is another decomposition and that p' is one of the primes occurring therein. From the preceding conclusion, we obtain $p' = p_\ell$, for some ℓ. Hence we deduce that, if the standard factorization for $\frac{n}{p'}$ is unique, then so also is that for n.

  The conclusion now follows by induction.

Greatest Common Divisor and Prime Decomposition

- It is simple to express the greatest common divisor (a, b) of elements a, b of $\mathbb{N}$ in terms of the primes occurring in their decompositions.

  We can write $a = p_1^{\alpha_1} \cdots p_k^{\alpha_k}$ and $b = p_1^{\beta_1} \cdots p_k^{\beta_k}$, where $p_1, \ldots, p_k$ are distinct primes and the α's and β's are non-negative integers.

  Then $(a, b) = p_1^{\gamma_1} \cdots p_k^{\gamma_k}$, where $\gamma_\ell = \min(\alpha_\ell, \beta_\ell)$.
- With the same notation, the **lowest common multiple** of a, b is defined by
$$\{a, b\} = p_1^{\delta_1} \cdots p_k^{\delta_k},$$
  where $\delta_\ell = \max(\alpha_\ell, \beta_\ell)$.
- Then we have
$$(a, b)\{a, b\} = ab.$$

Subsection 6

Properties of the Primes

Infinitude of Primes

Theorem

There exist infinitely many primes.

- Assume there are only finitely many, say n, different primes $p_1, p_2, \ldots, p_n$.

  The number $k = p_1 p_2 \cdots p_n + 1$ is not a prime, since it is greater than all available primes.

  So it has at least one prime factor, say p_m,

  i.e., there exists a number ℓ, such that $k = p_m \ell$.

  But now we get

$$\begin{array}{rcl} 1 & = & k - p_1 p_2 \cdots p_n = p_m \ell - p_1 p_2 \cdots p_n \\ & = & p_m(\ell - p_1 \cdots p_{m-1} p_{m+1} \cdots p_n), \end{array}$$

  i.e., $p_m \mid 1$, a contradiction.

One Consequence

Corollary

If p_n is the n-th prime in ascending order of magnitude, then p_m divides $p_1 \cdots p_n + 1$, for some $m \geq n+1$.

- The preceding proof showed that none of the primes $p_1, p_2, \ldots, p_n$ divides $p_1 p_2 \cdots p_n + 1$.

  It then follows that some prime $p_m > p_1, p_2, \ldots, p_n$ (i.e., such that $m \geq n+1$) must divide $p_1 p_2 \cdots p_n + 1$.

Bound on the Size of the n-th Prime

Theorem (Bound on the Size of p_n)

If p_n is the n-th prime in ascending order of magnitude, then $p_n < 2^{2^n}$.

- By induction on n.
  - For $n = 1$, $p_1 = 2 < 4 = 2^{2^1}$.
  - Suppose $p_k < 2^{2^k}$, for all $k \leq n$, where $n \geq 2$.
  - Then we obtain

$$
\begin{aligned}
p_{n+1} &\leq p_1 \cdots p_n + 1 < 2^2 \cdot 2^{2^2} \cdot 2^{2^3} \cdots 2^{2^n} + 1 \\
&= 2^{2+2^2+\cdots+2^n} + 1 = 2^{2(1+2+\cdots+2^{n-1})} + 1 \\
&= 2^{2\frac{2^n-1}{2-1}} + 1 = 2^{2^{n+1}-2} + 1 \\
&< 2^{2^{n+1}-1} < 2^{2^{n+1}}.
\end{aligned}
$$

The Prime Number Theorem

- Hadamard and de la Vallée Poussin proved independently in 1896 that, as $n \to \infty$,
$$p_n \sim n \log n,$$
where $f \sim g$ means $\frac{f}{g} \xrightarrow{n \to \infty} 1$.
- The result is equivalent to the assertion that the number $\pi(x)$ of primes $p \leq x$ satisfies
$$\pi(x) \sim \frac{x}{\log x}, \text{ as } x \to \infty.$$

Two Unsolved Problems

- **Goldbach Conjecture** (letter to Euler of 1742): Every even integer (>2) is the sum of two primes.
- **Twin-Prime Conjecture**: There exist infinitely many pairs of primes, such as $3,5$ and $17,19$, that differ by 2.
- By ingenious work on **sieve methods**, Chen showed in 1974 that these conjectures are valid if one of the primes is replaced by a number with at most two prime factors (assuming, in the Goldbach case, that the even integer is sufficiently large).
- Studies on Goldbach's conjecture gave rise to:
  - the **Hardy-Littlewood Circle Method** of analysis;
  - the celebrated **Theorem of Vinogradov**: Every sufficiently large odd integer is the sum of three primes.
- Recently, Yitang Zhang (UC-Santa Barbara), James Maynard (Oxford) and Terence Tao (UCLA) have contributed to new breakthroughs towards proving the Twin-Prime Conjecture.