Chapter 3

Introduction to Mathematical Systems (Abridged)

3.1 Basic Rationale for Axioms and Introduction

The theories of arithmetic, geometry, logic, sets, calculus, analysis, algebra, number theory, etc. were developed by many different mathematicians over centuries, but reached a rigorous level by the nineteenth and early twentieth centuries. A group of mathematicians at the University of Göttengen in Germany undertook the task of attempting to reduce the concepts in various fields to the simplest statements they could possibly assume such that the theory of a designated area must follow from said assumptions.

They took their example from Euclid and his series of books, *Elements*, in which Euclid proposed or collated proposals such that certain assumptions were made about geometry which all other facts about geometry could be deduced from those assumptions. Indeed, if memory serves me correctly (I will check and not rely solely on my rather imperfect memory) Euclid's fifth postulate was proposed but was not stated such that it was necessarily a postulate. He included it in the list so that it could be considered and left it to others to determine if indeed it was a postulate (something that had to be assumed) or followed from the other postulates. The fifth postulate was that for every line l and every point p not lying on l there exists a line m containing p such that m is parallel to l.[1]

The Germans were led by the mathematician, David Hilbert, and worked fervently on questions of axiomatics (as well as other things). They were known as reductionists for they were attempting to reduce the mathematical claims of the time to rigorous systems.

[1]Actually Euclid's fifth postulate was stated differently. The statement given here is the logically equivalent version of Euclid's fifth postulate from John Playfair's book, *Elements of Geometry*, published in 1795 or so.

Another great mathematician, George Boole, also worked on such questions. It is not to be assumed that such investigations met with universal acclaim, indeed Bertrand Russell and others thought the exercise rather ridiculous. They thought there was a need to study logic in its pure form and forgo such endeavours. Other mathematicians and logicians thought both the school to which we consider Russell a member, the logicians, and Hilbert, the axiomaticians, were wrong and that mathematics should be considered from an intuitive perspective. Such intuitionists as Kurt Gödel viewed mathematics as art. Thus, the discussion of the basic rationale for axiom systems does not imply that there is but one way to view mathematics, but we are in this course considering and examining mathematics from an axiomatic perspective.

The axiomatic approach to sets, logic, analysis, etc. is one of the most impressive accomplishments of modern mathematics. Concepts which were vague or indistinct took on the property of clarity. Precise meanings replaced quasi-definitions. Adequate axioms established the foundation of modern mathematics for they provided clear, unambiguous, and understandable premises for theories which before were sound but people *didn't know why they were sound*.

This is not to say that the original work of mathematicians was always without its problems. Around the turn of the twentieth century, Bertrand Russell proposed a construction which allowed for a paradox to be derived from the axioms system of the day proposed for set theory. There was an axiom, called the 'axiom' of abstraction, proposed by Frege which stated that given any property there exists a set whose members are those entities. Russell defines a set such that it was the set of all things which have the property of not being members of themselves. Suppose there was **a** universal 'set.' Let us call it U.

Let us call Russell's 'set' R. Now, what is R? $R = \{S \in U : S \notin S\}$. Now, the rub!

Does R belong to itself?

Suppose $R \in R$. Then since $R \in R$, it ***must*** be the case that $R \notin R$! Which is (of course) a contradiction.

Suppose $R \notin R$. Since $R \notin R$, it ***must*** be the case that $R \in R$! Which is (of course) a contradiction.

So, the absolutely hideous situation exists such that $R \in R \Leftrightarrow R \notin R$! So, the 'set' the set of all sets (the idea of **a** universal 'set') cannot exist. This of course, implies that the 'axiom' of abstraction is not an axiom of set theory (unfortunately for Frege, but fortunately for us).

This should illustrate for you, the reader, that one can call something a rule, a definition, an axiom, a postulate, a theorem, a lemma, or a corollary *but that does not mean it is*. It can only be such under the conditions that the axioms must be consistent (e.g.: they do not contradict each other) and all else follows from the axioms.

Nonetheless, it has been shown in the twentieth century that no system is compleat. So, one must pay careful attention to the axioms. Careless claims made that are intuitively appealing but are not derived from the axioms cannot be allowed.

Now recall that with a logic claim from chapter one that we proved we began by declaring what we were assuming (the premises) and deduced that the conclusion followed from those premises. We may have used a direct argument or indirect; but the important point is that the conclusion was derived from the premises. Note the premises did not contradict each other (so they were consistent) for if any subset of the set of premises were inconsistent, then we could not deduce a conclusion (since $F \Rightarrow T$ is true and $F \Rightarrow F$ is true).

Likewise, each branch of mathematics starts with a set of premises - - assumptions that are to be agreed are going to be assumed. These premises are the postulates or axioms.[2] Statements deduced from these assumptions are lemmas, theorems, or corollaries while the processes of deduction leading to these statements are the proofs themselves. Examples, definitions, and illustrations adjoin the lemmas, theorems, and corollaries in order to illuminate the concepts, to illustrate the principle, or to create new ideas.

The basic rules of language that is employed is the syntax whilst the meaning assigned to the symbols, words, etc. are the semantics of the language. We have already introduced much of the mathematical syntax in chapter one and two and we have introduced many of the semantics of logic and set theory in those chapters. Each branch of mathematics has its peculiarities (I warn you) - - so there is not necessarily a semantic standard. For example, in real analysis a function is defined from a set to another such that the first set is termed the domain and the second set is termed the codomain. The subset of the codomain that has associated with it at least one element of the domain is called the range. However, in probability theory, the first set for a probability function is termed the range. This should illustrate for you that whenever reading a math text find the glossary, index, and list of abbreviations and mark them! What one may think is *the* standard use of a symbol is not - - there really is not a standard.

Before each area of mathematics is discussed, acceptable syntactic and semantic rules must be adopted and one must understand the syntax and semantics in order to have any hope of understanding the area. When one notes that $|\mathbb{R}| = |\mathbb{I}| = \aleph_1$ and that $|\mathbb{Q}| = |\mathbb{N}| = \aleph_0$; that $\aleph_0 + 1 = \aleph_0$; whilst $\aleph_1 > \aleph_0$; hence, $|\mathbb{I}| > |\mathbb{Q}|$ one needs an understanding of relations, sets, functions, and cardinality in order to understand

[2]Axiom: from the Greek $\alpha\xi\iota\omega\mu\alpha$ loosely translated to mean that which is self-evident or thought to be fitting. Some thought axioms were self evident; but twentieth century logicians showed that self-evident is a rather dangerous concept. Hence, we shall adopt the position that the axioms are those primitive statements that are generally agreed to and that when we are going to study a particular branch of mathematics must be adhere to or obeyed.

transfinite arithmetic. Indeed, if one were to switch to ordinality much changes for $\omega_0 + 1 > \omega_0$ but $1 + \omega_0 = \omega_0$ one can be truly confused by this (and hopefully you are; for if you are not then you should not be in this class - for you know too much to waste time in this introductory level mathematics class). Hopefully this illustrates the idea that there is a language that is mathematics and that it is an exciting field with truly remarkable ideas that one can learn and master; but that it is a building process that leads us to these really astonishingly beautiful ideas.

One cannot run before one walks and one cannot walk before one crawls. You, the student, have passed the crawling stage and are in the walking stage; so please do not be impatient and imprudent and try to run before walking; but I digress.

The first component of many axiom systems is the notion of atoms.

Definition 3.1.1. Atoms (or primitive statements) are the undefined terms that are agreed to. For example in Euclidean geometry these atoms would be: there exist a point, a line, and a plane. A point has no dimension, a line has one dimension, and a plane has two dimensions. You should note that these atoms are used in many fields besides geometry and are some of the basic building blocks of mathematical thought.

A set of axioms is **consistent** if and only if it is impossible to deduce a statement and it logical opposite (e.g.: an axiom system is consistent $\Leftrightarrow$ one cannot deduced $P \wedge \neg P$ for some statement P). For us, this is an important point because if a set of axioms is inconsistent, then it is of no use.

Finally an axiom system needs a set of **rules of inference** so that theorems, examples, counter-examples, lemmas, and corollaries may be deduced from the axioms and definitions may be created to define terms, ideas, etc.

Thankfully, for our studies the basic rules of inference of propositional and syllogistic calculus (logic) are the rules used in classical mathematics.

In principle any set of consistent axioms can be studied, but the choice of axioms that we oft study (especially in a rigorous undergraduate mathematics curriculum) is not chosen in a capricious manner. Recall the discussion of the Göttengen mathematicians, they were central to much of modern mathematics because they attempted to reduce down to the axiomatic level the practical, useful, applied, and pure mathematics of the day in order to better understand that which was being claimed was true or false and to find the justification for things like calculus, topology, algebra, etc.

Mathematical theories are now understood to follow from axiom systems so that by deductive reasoning the theorems, examples, counter-examples, lemmas, and corollaries can be proven based on said axioms. Definitions are presented to clarify ideas, terms, etc. Consider the foundation of the system is the set of axioms; thus, the theory literally and figuratively is built upon that foundation. If the foundation be shoddy, then the theory collapses. So there is great import in ensuring the system is consistent, the syntactic rules

are sound, and the rules of inference are understood and properly executed, so that proofs and counterexamples are not possibly correct, maybe so, or any other nonsensical relativistic term but are declaratively true or false.

Example 3.1.1. *Let $U = \{\alpha, \beta, \gamma, \delta, \varepsilon, \omega\}$. Let $A = \{\alpha, \beta, \gamma, \delta\}$.*

Let '$\#$' be the operator such that an element of A '$\#$' another element of A is defined by the table below. The operation is pound; so we 'pound' two elements together. The elements of the universe are alpha, beta, gamma, delta, epsilon, and omega respectively. Hence, the reader can determine the elements of A. The operation pound is called a binary operation since it associates pairs of elements of A.

The elements are 'pounded' by reading the first column as the first element and the pound at the top of the first column and in the first row then reading the second element as the entry in the first row then follow the specified row and column to see what the elements 'pounded' together results.

$\#$	$\boldsymbol{\alpha}$	$\boldsymbol{\beta}$	$\boldsymbol{\gamma}$	$\boldsymbol{\delta}$
$\boldsymbol{\alpha}$	α	β	γ	δ
$\boldsymbol{\beta}$	β	γ	δ	α
$\boldsymbol{\gamma}$	γ	δ	α	β
$\boldsymbol{\delta}$	δ	α	β	γ

Table 3.1: "Pound" on A

Note that we did not define pound for all elements of the universe. Hence, there is no understanding as to what pound might do with an element of A and A^c or two elements of A^c.

Note further that $\alpha\#\beta$ is β, $\beta\#\alpha$ is β, $\gamma\#\delta$ is β, and $\delta\#\gamma$ is β. Let us examine this rather rudimentary mathematical system for its basic properties.

Note that whenever two elements are pounded together in A the result is a unique element in A. So, there is one and only one result when two elements of A are pounded together. So, this is an algebraic concept know as **closure**. The set A is **closed** under the binary operation pound. Inspection of the table suffices to *prove* that this claim is true. If one were to prove the claim that A is closed under pound, then a method of proof that will suffice is the **method of exhaustion**. This is because A is a finite set so that we need only make 16 observations. Please note that the method of exhaustion is *not* a valid method to prove that $\mathbb{N}$ is closed under '+' in ordinary arithmetic since $\mathbb{N}$ is an infinite set (more on this later).

Definition 3.1.2. A system $(S, \oplus)$ is **closed** if and only if given the binary operation $\oplus$ together with a pair of elements of S associates a unique element of S with the that pair of elements.

Note that when two elements, x and y, of A are pounded together in example 3.1.1 the result is the same as when y and x are pounded together. Once again we can use the method of exhaustion to prove this claim. Note that we used the variables x and y to denote arbitrary elements of the set A rather than any of the specific symbols α, β, γ, or δ. This is because we are trying to discuss the *general* truth that when two elements of A are pounded together the result is the same as when they are pounded in opposite order rather than a *specific* example of such like, beta pound gamma is delta and gamma pound beta is delta. So we say that $\#$ is a **commutative** or **abelian** operation on A when $x\#y = y\#x \ \ \forall x, y \in A$.

Definition 3.1.3. An operation $\otimes$ on a set S (for the system $(S, \otimes)$ is **commutative** if and only if given the binary operation $\otimes$ together with a pair of elements of S the order of the pair of elements does not matter, that is to say that element one operated $\otimes$ with element two is the same as element two operated $\otimes$ with element one.

Note further that when three elements, x, y, and z, of A are pounded together the result is the same no matter the order. That is to say that $x\#y\#z = (x\#y)\#z = x\#(y\#z)$ no matter what x, y, and z are in A. Once again we can use the method of exhaustion to prove this claim. Note that we used the variables x, y, and z to denote arbitrary elements of the set A rather than any of the specific symbols α, β, γ, or δ (even though there are but four elements in A). So we say that $\#$ is an **associative** operation on A when $x\#y\#z = (x\#y)\#z = x\#(y\#z) \ \ \forall x, y, z \in A$.

Definition 3.1.4. An operation $\odot$ on a set S (for the system $(S, \odot)$ is **associative** if and only if given the binary operation $\odot$ together with any elements of S the order of the execution of the operation does not matter, that is to say that element one operated $\odot$ with element two then that result operated $\odot$ with element three is the same as element two operated $\odot$ with element three first, the result of which when operated $\odot$ with element one yields the same result.

Note another interesting property that $(A, \#)$ exhibits. There is an element such that that element pounded with any element yields the element. That unique element is α. Notice $\alpha\#\beta = \beta, \alpha\#\delta = \delta$, and $\alpha\#\alpha$ is α (do not forget that it must be true for itself). We say that α is the **identity** element of A for the operation $\#$ on A. For a general definition consider:

Definition 3.1.5. An element x of the set S with the operation $*$ on S (for the system $(S, *)$ is called the **identity element** if and only if given the binary operation $*$ together with any element y in S it is the case that $x * y = y * x = y$.

Continuing by looking back to the example $(A, \#)$ we only needed to check the left operation (such as $\alpha\#\beta$ is β) since we already noted that $\#$ was commutative on A. Indeed note that $\alpha\#x$ is x and $x\#\alpha$ is $x \ \ \forall x \in A$.

So, the identity is a *particular* element that operates on every element in the set such that the operation with it 'changes nothing.'

Finally let us consider that there are other interesting elements in the system $(A, \#)$. That is easily proven using the method of exhaustion since $\alpha\#\beta$ is β, $\gamma\#\gamma = \alpha$, and $\alpha\#\alpha = \alpha$. We say that the elements x and y are **inverses** with respect to $\#$ when $x\#y = y\#x = i$ (the identity element). For a general definition consider:

Definition 3.1.6. The elements, p and q of the set S are **inverse elements** of each other with respect to the operation $\oslash$ on a set S (for the system $(S, \oslash)$ provided $p, q \in S$ it is the case that $p \oslash q = q \oslash p = i$.

Note for $(A, \#)$ we only needed to check the left operation (such as $\delta\#\beta$ is α) since we already noted that $\#$ was a commutative operation on A. Further notice that the definition of inverse elements was contingent on there being an identity. So, the definition of inverse elements would make no sense if there was not an identity.

Example 3.1.2. *Let $U = \mathbb{Z}$. Let us consider $\mathbb{N}$. Let '-' be the operator such that it is standard subtraction. So we define $x - y$ for all x and y in $\mathbb{N}$ to be the normal difference between to natural numbers. Note that $(\mathbb{N}, -)$ is not closed since $3 \in \mathbb{N}$ but $3 - 3 = 0$ and $0 \notin \mathbb{N}$. The other types of definitions for operations or elements for $(\mathbb{N}, -)$ are not operative. However, for the sake of understanding, please consider $(\mathbb{N}, -)$ does not have an identity. So, inverses under subtraction are out of the question. Note that '-' is not a commutative binary operation for $\mathbb{N}$. Also, notice that '-' is not an associative binary operation for $\mathbb{N}$ since $13 - (2 - 5)$ does not exist for $(\mathbb{N}, -)$ but $(13 - 2) - 5$ is well-defined for the natural numbers and is 6.*

Example 3.1.3. *Let $U = \mathbb{Z}$. Let us consider $A = \mathbb{Z}$. Let '-' be the operator such that it is standard subtraction. So we define $x - y$ to be the normal difference between to integers. Note now $(\mathbb{Z}, -)$ is closed.*
We can now justifiably consider the other definitions on this system $(\mathbb{Z}, -)$. It has an identity, 0. Indeed, subtractive inverses exist. Note, however, that '-' is not a commutative binary operation for $\mathbb{Z}$ since $3 - 2 \neq 2 - 3$. Also, notice that '-' is not an associative binary operation for $\mathbb{Z}$ since $13 - (2 - 5) = 13 - (-3) = 16$ but $(13 - 2) - 5 = 6$.

So, when we compare and contrast examples 3.1.2 and 3.1.3 we see that the set can make an important contribution to the discussion of mathematical systems. So too can the operation for consider the following example:

Example 3.1.4. *Let $U = \mathbb{Z}$. Let us consider $\mathbb{Z}$. Let '+' be the operator such that it is standard addition. So we define $x + y$ to be the normal sum of integers for any $x \in \mathbb{Z}$ and $y \in \mathbb{Z}$. Note that $(\mathbb{Z}, +)$ is closed. Note $(\mathbb{Z}, +)$ has an identity, 0. Indeed, additive inverses exist. Note that $+$ is a commutative binary operation for $\mathbb{Z}$. Also, note $+$ is an associative binary operation for $\mathbb{Z}$.*

So, when we compare and contrast examples 3.1.2, 3.1.3, and 3.1.4 we see that not only the set but the operation is very important to consider. Slight changes in either the operation or the set can cause each of the five properties discussed to be true or false, but not both. Note that for clarity we used the vernacular, 'true or false but not both' to properly represent $\vee$.

There are other definitions that are generalisations of the standard properties of the real numbers that can be noted. However, considering just the five properties here gives you, the student, the exposure to and the experience with abstract mathematical systems that is needed at this stage of your development. This is not to say that you, the student, cannot delve further into this topic if you are interested; it is merely to say that the subject will be expanded to include other operations, sets, and mathematical systems as your mathematics studies progress (or don't and this is the only college mathematics course you take).

3.1.1 Exercises

Exercise 3.1.1. Let $U = \mathbb{R}$ and let $A = \mathbb{N}$. Let the binary operation be the usual multiplication, $\cdot$.

1. Is $(\mathbb{N}, \cdot)$ closed? If so, why. If not show an example why it is not.
2. For $(\mathbb{N}, \cdot)$ does $\cdot$ fulfil the conditions to be commutative? If so, why. If not show an example why it is not.
3. For $(\mathbb{N}, \cdot)$ does $\cdot$ fulfil the conditions to be associative? If so, why. If not show an example why it is not.
4. For $(\mathbb{N}, \cdot)$ does there exist an identity element? If so, why. If not show an example why it is not.
5. For $(\mathbb{N}, \cdot)$ does every element have an inverse? If so, why. If not show an example why it is not.

Exercise 3.1.2. Let $U = \mathbb{R}$ and let $A = \mathbb{R}$. Let the binary operation be the usual addition, $+$.

1. Is $(\mathbb{R}, +)$ closed? If so, why. If not show an example why it is not.
2. For $(\mathbb{R}, +)$ does $+$ fulfil the conditions to be commutative? If so, why. If not show an example why it is not.
3. For $(\mathbb{R}, +)$ does $+$ fulfil the conditions to be associate? If so, why. If not show an example why it is not.

4. For $(\mathbb{R}, +)$ does there exist an identity element? If so, why. If not show an example why it is not.

5. For $(\mathbb{R}, +)$ does every element have an inverse? If so, why. If not show an example why it is not.

Exercise 3.1.3. Let $U = \mathbb{R}$ and let $A = \mathbb{R}$. Let the binary operation be the usual multiplication, $\cdot$.

1. Is $(\mathbb{R}, \cdot)$ closed? If so, why. If not show an example why it is not.

2. For $(\mathbb{R}, \cdot)$ does $\cdot$ fulfil the conditions to be commutative? If so, why. If not show an example why it is not.

3. For $(\mathbb{R}, \cdot)$ does $\cdot$ fulfil the conditions to be associate? If so, why. If not show an example why it is not.

4. For $(\mathbb{R}, \cdot)$ does there exist an identity element? If so, why. If not show an example why it is not.

5. For $(\mathbb{R}, \cdot)$ does every element have an inverse? If so, why. If not show an example why it is not.

Exercise 3.1.4. Let $U = M$ such that $M = \{\bigcirc, \triangle, \square\}$. Define the binary operation $*$ on M by the following table The elements are 'starred' by reading the first column as the first element and the 'star' at the top of the first column and in the first row then reading the second element as the entry in the first row then follow the specified row and column to see what the elements 'starred' together results.

$*$	$\triangle$	$\bigcirc$	$\square$
$\triangle$	$\bigcirc$	$\triangle$	$\square$
$\square$	$\triangle$	$\square$	$\bigcirc$
$\bigcirc$	$\square$	$\bigcirc$	$\triangle$

Table 3.2: "Asterisk" on M

1. Find $\triangle * \triangle$

2. Find $\triangle * \square$

3. Find $\square * \triangle$

4. Find $\bigcirc * \triangle$

5. Is it the case that $\bigcirc * \square = \square * \bigcirc$?
6. Is it the case that $(\triangle * \bigcirc) * \square = \triangle * (\bigcirc * \square)$?
7. Is $(M, *)$ closed?
8. Is $*$ commutative on M?
9. Is $*$ associative on M?
10. Does there exist an identity under $*$ on M? If so, what is it?
11. Does there exist an inverse for each element under $*$ on M? If so, what is the inverse for $\bigcirc$? If so, what is the inverse for $\triangle$? If so, what is the inverse for $\square$?

Exercise 3.1.5. Let $U = M$ such that $M = \{\bigcirc, \triangle, \square\}$. Define the binary operation $\perp$ on M by the following table The elements are 'botted' by reading the first column as the first element and the 'bot' at the top of the first column and in the first row then reading the second element as the entry in the first row then follow the specified row and column to see what the elements 'botted' together results.

$\perp$	$\triangle$	$\bigcirc$	$\square$
$\triangle$	$\triangle$	$\triangle$	$\square$
$\square$	$\square$	$\square$	$\bigcirc$
$\bigcirc$	$\square$	$\bigcirc$	$\bigcirc$

Table 3.3: "Bot" on M

1. Find $\bigcirc \perp \bigcirc$
2. Find $\triangle \perp \square$
3. Find $\square \perp \triangle$
4. Find $\bigcirc \perp \triangle$
5. Is it the case that $\bigcirc \perp \square = \square \perp \bigcirc$?
6. Is it the case that $(\triangle \perp \bigcirc) \perp \square = \triangle \perp (\bigcirc \perp \square)$?
7. Is $(M, \perp)$ closed?
8. Is $\perp$ commutative on M?

9. Is $\perp$ associative on M?
10. Does there exist an identity under $\perp$ on M? If so, what is it?
11. Does there exist an inverse for each element under $\perp$ on M? If so, what is the inverse for $\bigcirc$? If so, what is the inverse for $\triangle$? If so, what is the inverse for $\square$?

Exercise 3.1.6. Let $U = \mathbb{N}$ such that $A = \{1, 2, 3\}$. Define the binary operation $\sqcap$ on A by the following table The elements are 'capped' by reading the first column as the first element and the 'cap' at the top of the first column and in the first row then reading the second element as the entry in the first row then follow the specified row and column to see what the elements 'capped' together results.

$\sqcap$	**2**	**1**	**3**
2	1	2	3
3	2	3	1
1	3	1	2

Table 3.4: "Cap" on A

1. Find $2 \sqcap 2$
2. Find $2 \sqcap 3$
3. Find $3 \sqcap 2$
4. Find $1 \sqcap 2$
5. Is it the case that $1 \sqcap 3 = 3 \sqcap 1$?
6. Is it the case that $(2 \sqcap 1) \sqcap 3 = 2 \sqcap (1 \sqcap 3)$?
7. Is $(A, \sqcap)$ closed?
8. Is $\sqcap$ commutative on A?
9. Is $\sqcap$ associative on A?
10. Does there exist an identity under $\sqcap$ on A? If so, what is it?
11. Does there exist an inverse for each element under $\sqcap$ on A? If so, what is the inverse for 1? If so, what is the inverse for 2? If so, what is the inverse for 3?
12. Look at the previous exercises. Do you notice something similar?

3.2 A Different Base

Why do we count as we do? What is the reason? We use the base ten system and we define digits to be the universe $D = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$. So, our digits that are from the set natural numbers star ($\mathbb{N}^* = \{0, 1, 2, 3, 4, 5, ...\}$) simply show positional meaning to the powers of ten. Hopefully, we all recall that the number $1,237$ simply means that we have one 10^3, two 10^2, three 10^1, and seven 10^0. So our positional method (the Hindu-Arabic numeration system so named based on the symbols being developed in India then transmitted through the Muslim caliphate to Africa and Europe) is an elegant and useful method for expressing the natural numbers (with zero) and is extended to the integers, rationals, etc.

We were probably told that the Hindu-Arabic numeration system is that which it is and it is the best way to do it. But in the course of our existence we use other systems; for example, modular or base 12 and base 60 (sort of) for time, base 12 and base 3 (sort of) for English measurement of distance, etc. It **is** the best system; for if one studies other numeration systems one would find they lack the ease of operation and the rigour of the Hindu-Arabic numeration system. You are free to study other systems and compare them to this one. Our discussion will centre on base and modular arithmetic using the Hindu-Arabic numeration system or an extension of them.

Formally, a natural number can be expressed in any base system such that the base is well defined so that each position represents groups of powers of the base.

Consider $1,237$ in base ten means one 10^3, two 10^2, three 10^1, and seven 10^0; but, if this were base nine then $1,237$ would mean one 9^3, two 9^2, three 9^1, and seven 9^0. This is meaningful, but for base five, for example it would not because since $7 > 5$ how could one have digit of 7 in base 5? One can't; so, the digits for each type of base depend on the base.

For base 2 the set of digits is $T = \{0, 1\}$;
for base 3 the set of digits is $H = \{0, 1, 2\}$;
for base 4 the set of digits is $F = \{0, 1, 2, 3\}$;
for base 5 the set of digits is $V = \{0, 1, 2, 3, 4\}$;
for base 6 the set of digits is $X = \{0, 1, 2, 3, 4, 5\}$;
for base 7 the set of digits is $S = \{0, 1, 2, 3, 4, 5, 6\}$;
for base 8 the set of digits is $E = \{0, 1, 2, 4, 5, 6, 7\}$; and,
for base 9 the set of digits is $N = \{0, 1, 2, 4, 5, 6, 7, 8\}$.
One can extend past base 10 in more than one way. Let us use the 'alpha-numeral' system such that for base eleven (we can't say '11' since that is base 10 the set of digits is $L = \{0, 1, 2, 4, 5, 6, 7, 8, 9, T\}$; for base twelve the set of digits is
$W = \{0, 1, 2, 4, 5, 6, 7, 8, 9, T, E\}$; etc.

Now in each system the value in each position represents the number of powers of the base from right to left. Hence, the number 11010 in base 2 is well defined and for clarity

when we are referencing a number in an alternate base system let us use a subscript to clarify the meaning; so, let us let 'one one zero one zero base two,' be written as 11010_2.

It is 1 of 2^4, 1 of 2^3, 0 of 2^2, 1 of 2^1 , 0 of 2^0. Hence it is 1×2^4, 1×2^3, 0×2^2, 1×2^1 , and 0×2^0 using the elementary sign for multiplication.

So, it is $(1 \cdot 2^4) + (1 \cdot 2^3) + (0 \cdot 2^2) + (1 \cdot 2^1) + (0 \cdot 2^0)$.

Hopefully, it is facile to see that we therefore have $(16) + (8) + (0) + (2) + (0)$.

So, $11010_2 \equiv 26_{10}$ in standard decimal (base ten) form; we shall use the symbol for logically equivalent ($\equiv$) since that is what we are expressing that the two concepts are indeed the same only they are expressed in different systems.

Example 3.2.1. *Suppose we are presented with* 113241_5. *What is it?*

Clearly it is $(1 \cdot 5^5) + (1 \cdot 5^4) + (3 \cdot 5^3) + (2 \cdot 5^2) + (4 \cdot 5^1) + (1 \cdot 5^0)$*; which is by the axioms of the natural numbers,* $(1 \cdot 5^0) + (4 \cdot 5^1) + (2 \cdot 5^2) + (3 \cdot 5^3) + (1 \cdot 5^4) + (1 \cdot 5^5)$
$= (1 \cdot 1) + (4 \cdot 5) + (2 \cdot 25) + (3 \cdot 125) + (1 \cdot 625) + (1 \cdot 3,125)$.
$= (1) + (20) + (50) + (375) + (625) + (3,125) = 4,196$.

Principle 3.2.1. *Suppose a is expressed in a base other than* 10. *To covert to base* 10 *one expands the number by definition of the powers of the base; executes the arithmetic (in base 10); and, then multiplies and finally adds (ask yourself why?).*

Now consider $3,401$. Suppose we wish to convert it to base 6 (yes, I am aware this is rather odd but bear with me). Since conversion from base 'A' to base ten was done through expansion with multiplication and addition it is logical to conclude that this process will require division and subtraction (explain why this seems reasonable to yourself). However, we need the powers of 6.
Note that $6^0 = 1, 6^1 = 6, 6^2 = 36, 6^3 = 216, 6^4 = 1,296, 6^5 = 7,776$, and so forth. We only need those powers less than or equal to $3,401$ since there can be no groups of size $7,776$ or more to allot.

Now, note the algorithm we shall use.

Note that $1,296 \cdot 2 = 2,592$; but, $1,292 \cdot 3 = 3,888$ (which is too much - $3,888$ goes into $3,401$ zero times - we are 'over').

$$\begin{array}{r} 2 \\ 1296\overline{)\,3401} \\ \underline{2592} \\ 809 \end{array}$$

$3,888 \div 1,296 = 2$ with remainder 809

$$\begin{array}{r} 3 \\ 216\overline{)\,809} \\ \underline{648} \\ 161 \end{array}$$

Hence, $216 \cdot 3 = 648$ with remainder 161

$$\begin{array}{r} 0 \\ 161\overline{)\,36} \end{array}$$

Hence, $36 \cdot 4 = 144$ with remainder 17

$$\begin{array}{r} 0 \\ 17\overline{)\,6} \end{array}$$

Therefore, $6 \cdot 2 = 12$ with remainder 5.

$$\begin{array}{r} 0 \\ 5\overline{)\,1} \end{array}$$

Finally, note that $5 \cdot 1 = 5$ with remainder 0 (leaving no remainder).
Hence, it is the case that 3401_{10} converted to base 6 proceeds as:
$3401 = 2596 + 648 + 144 + 12 + 5 =$
$(2 \cdot 1296) + (3 \cdot 216) + (4 \cdot 36) + (2 \cdot 6) + (5 \cdot 1) =$
$(2 \cdot 6^4) + (3 \cdot 6^3) + (4 \cdot 6^2) + (2 \cdot 6^1) + (5 \cdot 6^0) \Rightarrow$
$3401 \equiv 23425_6$.

Now, before proceeding any further explain what the algorithm was; how it was used; why it was used; and, opine as to its generalisation.

Example 3.2.2. *Consider* 711. *Let us convert this to base* 2. *You can create the algorithm yourself; suffice it to say the powers of two are*
$1, 2, 4, 8, 16, 32, 64, 128, 256, 512, 1,024, 2,048,$ *etc.*[3] *We only need begin our work with* 512. *I will do the algorithm in a slightly different way.*
$711 - 512 = 199$.
$199 - 128 = 81$.
$81 - 64 = 17$.
32 *into* 17 *yields* 0 *with remainder* 17 *so proceed to* 2^4.
$17 - 16 = 1$.
We divide by $8, 4,$ *and* 2 *and also get zeros.*
Finally 1 *divided by* 1 *is* 1 *with remainder zero. So we get* 111010001_2. *Thus,*
$711 \equiv 111010001_2$.
Notice that the base 10 *subscript is not written (it can be; but, does not have to be written– the other subscripts must be written for clarity, precision, and proper form).*

[3] We had to memorise such in Primary School back in the day (1960s).

Example 3.2.3. *Let us consider* 711 *again only this time let us convert it to base 12. The powers of twelve are* $1, 12, 144, 1, 728$, *etc.*

$$\begin{array}{r} 4 \\ 144\overline{)\,711} \\ \underline{576} \\ 135 \end{array}$$

Note that $144 \cdot 4 = 576$ *and we have a remainder of* 135.

$$\begin{array}{r} 11 \\ 12\overline{)\,135} \\ \underline{120} \\ 15 \\ \underline{12} \\ 3 \end{array}$$

Note that $12 \cdot 11 = 132$ *and we have a remainder of* 3.
But, we cannot use '11' so we use E.

$$\begin{array}{r} 3 \\ 1\overline{)\,3} \\ \underline{3} \\ 0 \end{array}$$

Now, Note that $1 \cdot 3 = 3$. *So, we are done. Hence,* $711 \equiv 4E3_{12}$.

Principle 3.2.2. *Suppose a is expressed in base* 10 *and we wish to covert to a base other than* 10. *To covert to base a one uses division and subtraction (noting the remainder); executes the arithmetic (in base 10); and, then expresses the results on the new base a.*

Now, the easiest way to convert from base A to base B (where A and B are not 10) is into convert to base ten then out of it.

Example 3.2.4. *For example, converting* 312_4 *to base* 7 *entails considering that*

312_4 *is* $(3 \cdot 4^2) + (1 \cdot 4^1) + (2 \cdot 4^0) =$
$(3 \cdot 16) + (1 \cdot 4) + (2 \cdot 1) =$
$48 + 4 + 2 = 54.$

Now the powers of seven are $1, 7, 49, 343, 2, 401, 16, 807$, *etc. We could begin with 343; but that would be foolish. We need only begin with* 49. *Note that* $49 \cdot 1 = 49$. *So, we subtract* 49 *from* 54 *leaving* 5 *as a remainder.*
Note that $7 \cdot 0 = 0$. *So, we don't subtract; still leaving 5 as a remainder.*
Now, note that $1 \cdot 5 = 5$. *Subtract and we have remainder* 0. *So, we are done.*
Hence, $312_4 \equiv 105_7$. *Please note that we are using transitivity to deduce this.*

There are short-cuts and other tricks that (long ago) we studied in Sister Rose Dominic's fifth grade class at St. Margaret's School; but, suffice it to say that is the gist of base systems. So, try some!

3.2.1 Exercises in Basal Arithmetic

Exercise 3.2.1. Convert the following numbers to base ten.

A. 813_9	G. 313_4	M. 6611_8
B. 110311_4	H. 3103013_4	N. 1166_8
C. 110101_2	I. 1010101_2	O. 6116_8
D. 03413_5	J. 34013_5	P. $3ET1_{12}$
E. 110311_7	K. 110131_7	Q. 1562_{12}
F. 1661_8	L. 6161_8	R. 111_{11}

Exercise 3.2.2. Convert 913 the following bases

A. base 2
B. base 3
C. base 4
D. base 5
E. base 6
F. base 8
G. base 9
H. base 12
I. base 16 (define the symbols used)
J. base 7

Exercise 3.2.3. Convert the following numbers to the specified base.

A. 813_9 to base 3	G. 313_4 to base 5	M. 661_8 to base 6
B. 110311_4 to base 2	H. 3103013_4 to base 9	N. 1166_8 to base 6
C. 110101_2 to base 3	I. 1010101_2 to base 8	O. 665_8 to base 6
D. 3413_5 to base 8	J. 34013_5 to base 7	P. $ET1_{12}$ to base 2
E. 1101_7 to base 12	K. 1101_7 to base 5	Q. 1562_{12} to base 5
F. 661_8 to base 7	L. 661_8 to base 4	R. 121_{11} to base 5

Exercise 3.2.4. Convert 10101001001_2 to base 10.

Exercise 3.2.5. Let the digits for base 16 be $\{0, 1, 2, 4, 5, 6, 7, 8, 9, \alpha, \beta, \gamma, \delta, \varepsilon, \omega\}$ where α is ten; β is eleven; γ is twelve; δ is thirteen; ε is fourteen; and, ω is fifteen. Convert the following to base 10:

A. 813_{16} B. $\beta\,\alpha\,\delta_{16}$ C. 256_{16} D. 101_{16} E. 1001_{16}

Exercise 3.2.6. Determine which is greater 1304_5 or 503_7 and justify your conclusion.

Definition 3.2.1. Define addition and multiplication in other bases to be as we would naturally assume them to be.
e.g.: $a_b + c_b = d_b$ if and only if $a_b \equiv x_{10}$ and $c_b \equiv y_{10}$ yields $x + y = z$ and $z \equiv d_b$.
$a_b \cdot c_b = f_b$ if and only if $a_b \equiv x_{10}$ and $c_b \equiv y_{10}$ yields $x \cdot y = w$ and $w \equiv f_b$.

Exercise 3.2.7. Compute the following:

A. $3_9 + 7_9$	G. $11_2 + 1_2$	M. $1_2 \cdot 1_2$
B. $131_4 + 1311_4$	H. $31_4 + 11_4$	N. $6_8 \cdot 7_8$
C. $110101_2 + 110101_2$	I. $31_4 + 32_4$	O. $3_8 \cdot 2_8$
D. $110101_2 + 1010_2$	J. $31_4 \cdot 11_4$	P. $3ET1_{12} + 7_{12}$
E. $11_2 \cdot 11_2$	K. $31_4 \cdot 3_4$	Q. $101_2 \cdot 11_2$
F. $813_9 + 712_9$	L. $110101_2 + 1111_2$	R. $101_2 \cdot 1001_2$

3.3 Modular Arithmetic

As discussed in the previous section, we count in other ways besides just base ten. The illustration that I alluded to was time, weight, distance, etc. that we in the United States follow rather than as most humans where they have acclimated to a decimal system for all but time.

Nonetheless, there is something really odd about our system. Think about the following:
Start with 1 minim. After a while we get 60 minims which is one dram. Later we finally have 8 drams which is an ounce. Sixteen ounces and we have a pint. Two pints and we have a quart. Four quarts and we have a gallon, and so for and so on. What of drams, ounces, pounds, tons, etc. not to mention inches, feet, yards, miles, etc.? Why measure time as 60 seconds for one minute, sixty minutes for an hour, but twenty-four hours for a day; and more preposterous months with 28, 29, 30, or 31 days but there are twelve of them so a year has 365 or 366 days?[4]

Well, it is convention and sometimes ten is not always the best under any and all circumstances, eh? So, what of these systems? They illustrate modular arithmetic (some with changes in the place values). Consider $11 + 5 = 4$. Note it is nonsense since the universe was not defined and it leads on to assume it is referencing real arithmetic. We are *not* in the business to play games and attempt to fool others, each other, or ourselves (though I am easily confused).

So, let us define modular arithmetic as follows.

Definition 3.3.1. Let $U = \mathbb{N}$, $p \in \mathbb{N}$, and $D = \{0, 1, 2, 3, ..., (p-1), p\}$. Let $a \in D$ and $b \in D$. Define $a +_p b$ to be $z \in D$ where z is the remainder from the division of p into $(a + b)$ under ordinary addition and division for the natural numbers.

So, for example let us see what "$+_4$" is (meaning $U = \mathbb{N}$, $4 \in \mathbb{N}$, and $D = \{0, 1, 2, 3\}$.
Let $0 +_4 0 = 0$, $0 +_4 1 = 1$, $0 +_4 2 = 2$, $0 +_4 3 = 3$,
$1 +_4 0 = 0$, $2 +_4 0 = 0$, $3 +_4 0 = 3$ all because 4 divided into any of these is zero with remainder whatever the non-zero constant was.

[4]Check this out: years ending with '00' but not divisible by 400 still only have 365 days. Ugh!

Consider $1 +_4 1 = 2$, $1 +_4 2 = 3$, $2 +_4 1 = 3$ for the same reasons.
So, we are left with the other possibilities:
$1 +_4 3 = 0$ since $4 \div 4 = 1$ with remainder 0.
$2 +_4 2 = 0$ since $4 \div 4 = 1$ with remainder 0.
$2 +_4 3 = 1$ since $5 \div 4 = 1$ with remainder 1.
$3 +_4 3 = 2$ since $6 \div 4 = 1$ with remainder 2.
$3 +_4 1 = 0$ since $4 \div 4 = 1$ with remainder 0.
$3 +_4 2 = 1$ since $5 \div 4 = 1$ with remainder 1.
Note that this is all good and well but it can be confusing. So, let us note the results in tabular form: This should look strikingly familiar (see section 3.1).

$+_4$	**0**	**1**	**2**	**3**
0	0	1	2	3
1	1	2	3	0
2	2	3	0	1
3	3	0	1	2

Table 3.5: Mod 4 Addition

You should be comfortable with defining the same for other mods (modular arithmetic operations with some natural number) such as: and: along with modular twelve

$+_5$	**0**	**1**	**2**	**3**	**4**
0	0	1	2	3	4
1	1	2	3	4	0
2	2	3	4	0	1
3	3	4	0	1	2
4	4	0	1	2	3

Table 3.6: Mod 5 Addition

$+_2$	**0**	**1**
0	0	1
1	1	0

Table 3.7: Mod 2 Addition (snore)

arithmetic (though technically T is ten and E is eleven (see a clock in hours).

We can also define modular multiplication:

Definition 3.3.2. Let $U = \mathbb{N}$, $p \in \mathbb{N}$, and $D = \{0, 1, 2, 3, ..., (p-1), p\}$. Let $a \in D$ and $b \in D$. Define $a \times_p b$ (or $a \cdot_p b$ which prints poorly so we won't use it) to be $w \in D$ where w is the remainder from the division of p into $(a \cdot b)$ under ordinary multiplication and division for the natural numbers.

Do the arithmetic in your head for mod 4 multiplication to get the following: Indeed

$\times_4$	**0**	**1**	**2**	**3**
0	0	0	0	0
1	0	1	2	3
2	0	2	0	2
3	0	3	2	1

Table 3.8: Mod 4 Multiplication

notice mod 6 multiplication:

$\times_6$	**0**	**1**	**2**	**3**	**4**	**5**
0	0	0	0	0	0	0
1	0	1	2	3	4	5
2	0	2	4	0	2	4
3	0	3	0	3	0	3
4	0	4	2	0	4	2
5	0	5	4	3	2	1

Table 3.9: Mod 6 Multiplication

Now reconsider $11 + 5 = 4$. Recall we are *not* in the business to play 'games.' Consider $+_{12}$. Note it is fine to consider $U = \mathbb{N}_{11}^*$ and let the set of digits, D be U.[5] $11 +_{12} 5 = 4$ is understandable from the applied problem of 11 a.m. and 5 hours is 4 p.m. (5 hours after 11 a.m. is 4 p.m.). We can deduce $9 +_{12} 4 = 1$ and we can realise $9 \times_{12} 4$ is meaningful as $9 +_{12} 9 +_{12} +9 +_{12} 9$ or simply $9 \times_{12} 4 = 0$ because $9 \times 4 = 36$ and $36 \div 6 = 6$ with remainder 0. Notice we do not use the universe $D = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, T, E\}$ as with base arithmetic. Also, note that commutativity and associativity hold. Now, does

[5]Recall, I suppose *technically* T should be the digit for ten and E for eleven; but, we will not do so for this example.

there exist a distributive property of modular multiplication over modular addition? If so, how would you argue it; if not, can you devise a counterexample?

It is perfectly acceptable for an application to define a system of modular arithmetic that includes $p \in \mathbb{N}$ for U as with time. Hence, in modular 12 arithmetic, $7 +_{12} 5 = 0$ but in practical applications most would say it is 12 since $12 \equiv 0$ in the sense of time (in hours). Once again defining the domain of discourse for the discussion (logic) and noting the universe (sets) is paramount to clearly and objectively investigating, understanding, and applying the mathematical system.

3.3.1 Exercises

Exercise 3.3.1. Construct the table for addition and multiplication mod 3.

Exercise 3.3.2. Construct the table for addition and multiplication mod 8.

Exercise 3.3.3. Calculate the following

A. $8 +_9 1$	B. $8 +_9 1 +_9 3$	C. $7 +_9 7 +_9 7$
D. $2 +_9 2 +_9 2$	E. $7 +_8 5 +_8 6$	F. $11 +_{12} 11 +_{12} 5$
G. $8 \times_9 1$	H. $5 \times_9 3$	I. $7 \times_9 7$
J. $6 +_8 5 +_8 7$	K. $1 +_2 1$	L. $1 +_2 1 +_2 1$
M. $7 \times_8 5 +_8 6$	N. $(7 \times_8 5) +_8 6$	O. $7 \times_8 (5 +_8 6)$
P. $0 \times_{45} 1$	Q. $5 \times_6 5 \times_6 5$	R. 7_6^3
S. 1_6^5	T. $5 \times_6 4 \times_6 3$	U. $3 \times_6 4 \times_6 5$
V. $4 \times_6 3 \times_6 5$	W. $5 \times_6 3 \times_6 4$	

X. $4 +_5 3 +_5 3 +_5 3 +_5 1 +_5 4 +_5 2$

Exercise 3.3.4. What is incorrect about the following statement? $9 \times_8 7 = 7 \times_8 9$.

Exercise 3.3.5. Compute the following (what do you realise?):

A. $8 -_9 7$	B. $7 -_9 8$	C. $1111 -_2 1111$
D. $4 -_5 2$	E. $4 -_5 3$	F. $4 -_5 4$
G. $3 -_5 2$	H. $3 -_5 3$	I. $3 -_5 4$
J. $2 -_5 2$	K. $2 -_5 3$	L. $2 -_5 4$

Exercise 3.3.6. Compare the results of this exercise set with exercise set for basal arithmetic. What patterns (if any) exist and can they be explained or an hypothesis formed as to the relationship between modular arithmetic and arithmetic of natural numbers in different bases?

Exercise 3.3.7. Can modular division be well defined? Why or why not?