Chapter 3
Introduction to Axioms, Mathematical Systems, Arithmetic, The Peano Axioms, and Mathematical Induction.

§ 3.1 BASIC RATIONALE FOR AXIOMS AND AN INTRODUCTION TO MATHEMATICAL SYSTEMS.

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. 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

[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.

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.

---

[2] Axiom: from the Greek αξιωμα 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.

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 $|\mathbb{C}| = |\mathbb{R}| = \aleph_1$ and that $|\mathbb{N}| = |\mathbb{Z}| = |\mathbb{Q}| = \aleph_0$; that $\aleph_0 + 1 = \aleph_0$; whilst $\aleph_1 > \aleph_0$; hence, $|\mathbb{I}| = \aleph_1$ one needs an understanding of relations, sets, functions, and cardinality in order to understand transfinite arithmetic. Indeed, if one were to switch to ordinality and not the transfinite ordinal equation $\omega_0 + 1 > \omega_0$ 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. **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 to us.

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 table 3.1.1. 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.

#	α	β	γ	δ
α	α	β	γ	δ
β	β	γ	δ	α
γ	γ	δ	α	β
δ	δ	α	β	γ

Table 3.3.1 (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 ($4 \cdot 4$ 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 3.1.1: A system $(S, \pitchfork)$ is **closed** if and only if given the binary operation $\pitchfork$ 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 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

$x \# y = y \# x \quad \forall\ x, y \in A.$

Definition 3.1.2: An operation ⫛ on a set S (for the system (S, ⫛)) is **commutative** if and only if given the binary operation ⫛ 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 (⫛) with element two is the same as element two operated (⫛) 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

$x \# y \# z = (x \# y) \# z = x \# (y \# z) \quad \forall\ x, y, z \in A.$

Definition 3.1.3: An operation ⫛ on a set S (for the system (S, ⫛)) is **associative** if and only if given the binary operation ⫛ together with any three elements of S the order of the execution of the operation does not matter, that is to say that element one operated (⫛) with element two then that result operated (⫛) with element three is the same as element two operated (⫛) with element three first, the result of which when operated (⫛) 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 3.1.4: 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 $x \fatsemi y = y \fatsemi x = y$.

Note for (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 $x \# y = y \# x = \alpha$ (the identity element). For a general definition consider:

Definition 3.1.5: An pair of elements x and y of the set S with the operation $\pitchfork$ on S (for the system (S, $\pitchfork$)) are called **inverse elements of each other with respect to** $\pitchfork$ if and only if given the binary operation $\pitchfork$ together with x and y it is the case that $x \pitchfork y = y \pitchfork x$ which results in the identity element.

Note for (A, #) we only needed to check the left operation (such as $\delta \# \beta$ is α) since we already noted that # was commutative 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 to be the normal difference between to natural numbers. Note that $(\mathbb{N}, -)$ is not closed since $3 \in \mathbb{N}$, but $3 - 3 = 0 \notin \mathbb{N}$. Note $(\mathbb{N}, -)$ 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 $\mathbb{Z}$. Let '–' be the operator such that it is standard subtraction. So we define x – y to be the normal difference between to integers. Note that $(\mathbb{Z}, -)$ is closed. Note $(\mathbb{Z}, -)$ has an identity, 0. Indeed, subtractive inverses exist. Note, however, that – is not a commutative binary operation for $\mathbb{Z}$. 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. 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 $\veebar$.

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.

§ 3.1 EXERCISES.

1. Prove that # from example 3.1.1 is commutative on A.
2. Prove that # from example 3.1.1 is associative on A.
3. Define the following on *U* from example 3.1.1:

#	α	β	γ	δ	ε	ω
α	α	β	γ	δ	ε	ω
β	β	γ	δ	α	ω	ε
γ	γ	δ	α	β	δ	β
δ	δ	α	β	γ	β	δ
ε	ε	ω	δ	β	α	γ
ω	ω	ε	β	δ	γ	α

A. Prove or disprove that *U* is closed under #.
B. Prove or disprove that # is commutative on *U*.
C. Prove or disprove that # is associative on *U*.
D. Prove or disprove that there exists an identity under # in *U*.
E. Prove or disprove that there exists an inverse for α in *U* under the operation #.
F. Prove or disprove that there exists an inverse for β in *U* under the operation #.
G. Prove or disprove that there exists an inverse for γ in *U* under the operation #.
H. Prove or disprove that there exists an inverse for δ in *U* under the operation #.
I. Prove or disprove that there exists an inverse for ε in *U* under the operation #.
J. Prove or disprove that there exists an inverse for ω in *U* under the operation #.

4. Let *U* = M such that M = {△, ▭, ○}. Define the binary operation Þ on M so that

Þ	△	▭	○
△	○	△	▭
▭	△	▭	○
○	▭	○	△

A. Find △ Þ △.
B. Find △ Þ ▭.
C. Find △ Þ ○.
D. Find ○ Þ ▭.
E. Find ○ Þ △.
F. Is there an identity element for Þ in M ?
G. Is it the case that (○ Þ △) Þ ▭ = ○ Þ (△ Þ ▭) ?
H. Is it the case that △ Þ ▭ = ▭ Þ △ ?
I. Is it the case that (△ Þ ▭) Þ ○ = △ Þ (▭ Þ ○) ?
K. Is it the case that (▭ Þ ○) Þ △ = ▭ Þ (○ Þ △) ?
L. Is there an inverse element for ○ in M under Þ ?
M. Is there an inverse element for ▭ in M under Þ ?

5. Let $U = \mathbb{Z}$. Let us consider $\mathbb{Z}$. Let '$\otimes$' be the operator such that it means, 'if the two integers are the same select the number and if the two numbers are not the same select the lesser of the two numbers.' So we define $x \otimes y$ to be x if $x = y$ [or y if one so desires], $x \otimes y$ to be x if $x < y$, and
$x \otimes y$ to be y if $x > y$.

A. Compute $5 \otimes 3$.

B. Compute $3 \otimes 3$.

C. Compute $5 \otimes 5$.

D. Compute $3 \otimes 3$.

E. Compute $(-5) \otimes 3$.

Justify.

F. Compute $5 \otimes (-3)$.

G. Compute $(-5) \otimes (-3)$.

H. Is $\mathbb{Z}$ closed under $\otimes$? Justify your response.

I. Does there exist an identity for $(\mathbb{Z}, \otimes)$? Justify.

J. Is $\otimes$ a commutative operation in $\mathbb{Z}$? Justify.

K. Is $\otimes$ an associative operation in $\mathbb{Z}$? Justify.

L. Does there exist an inverse for each $x \in \mathbb{Z}$ under $\otimes$?

6. Let $U = \mathbb{Z}$. Let us consider $\mathbb{N}$. Let '$\otimes$' be the operator such that it means, 'if the two natural numbers are the same select the number and if the two numbers are not the same select the lesser of the two numbers.' So we define $x \otimes y$ to be x if $x = y$ [or y if one so desires], $x \otimes y$ to be x if $x < y$, and $x \otimes y$ to be y if $x > y$.

A. Compute $5 \otimes 3$.

B. Compute $3 \otimes 3$.

C. Compute $51 \otimes 15$.

D. Compute $3 \otimes 3$.

E. Compute $1 \otimes 3$.

Justify.

F. Compute $5 \otimes 13$.

G. Compute $4701 \otimes 4701$.

H. Is $\mathbb{N}$ closed under $\otimes$? Justify your response.

I. Does there exist an identity for $(\mathbb{N}, \otimes)$? Justify.

J. Is $\otimes$ a commutative operation in $\mathbb{N}$? Justify.

K. Is $\otimes$ an associative operation in $\mathbb{N}$? Justify.

L. Does there exist an inverse for each $x \in \mathbb{N}$ under $\otimes$?

7. Let $U = \mathbb{Z}$. Let us consider $\mathbb{Z}$. Let '$\oplus$' be the operator such that it means, 'select the integer before the sign.' So we define x $\oplus$ y to be x.

A. Compute 5 $\oplus$ 3.

B. Compute 3 $\oplus$ 3.

C. Compute 5 $\oplus$ 5.

D. Compute 3 $\oplus$ 3.

E. Compute (-5) $\oplus$ 3.

F. Compute 5 $\oplus$ (-3).

G. Compute (-5) $\oplus$ (-3).

H. Is $\mathbb{Z}$ closed under $\oplus$? Justify your response.

I. Does there exist an identity for ($\mathbb{Z}$, $\oplus$)? Justify.

J. Is $\oplus$ a commutative operation in $\mathbb{Z}$? Justify.

K. Is $\oplus$ an associative operation in $\mathbb{Z}$? Justify.

L. Does there exist an inverse for each x $\in$ $\mathbb{Z}$ under $\oplus$? Justify.

6. Let $U = \mathbb{Z}$. Let us consider $\mathbb{N}$. Let '$\oplus$' be the operator such that it means, 'select the natural number before the sign.' So we define x $\oplus$ y to be x.

A. Compute 5 $\oplus$ 3.

B. Compute 3 $\oplus$ 3.

C. Compute 51 $\oplus$ 15.

D. Compute 3 $\oplus$ 3.

E. Compute 1 $\oplus$ 3.

F. Compute 15 $\oplus$ 3.

G. Compute 4301 $\oplus$ 4701.

H. Is $\mathbb{N}$ closed under $\oplus$? Justify your response.

I. Does there exist an identity for ($\mathbb{N}$, $\oplus$)? Justify.

J. Is $\oplus$ a commutative operation in $\mathbb{N}$? Justify.

K. Is $\oplus$ an associative operation in $\mathbb{N}$? Justify.

L. Does there exist an inverse for each x $\in$ $\mathbb{N}$ under $\oplus$? Justify.

§ 3.2 SOME FUNDAMENTAL AXIOM SYSTEMS.

The theories of arithmetic, geometry, logic, sets, calculus, analysis, algebra, etc. are all fundamentally supported by an axiom system. In some areas of mathematics, the axioms overlap to produce a tapestry of great detail that produces some very interesting results.

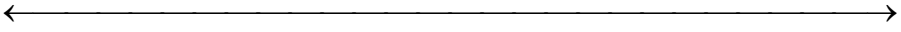
On the other hand, each individual area has its own nomenclature; definitions; theorems; etc. so caution is advised - - one should carefully review the foundations of a mathematical area and not assume that terminology that one is familiar with is used in the way in which he has become familiar.

A brief perusal of a statistics text would give a student a clear exposition of statistics, but he might be under a cloud of misunderstanding if he read the sentences: "Let X be a binomial random variable with parametres n and p such that $n \in \mathbb{N}$ and $p \in (0, 1)$. So, n defines the range of the probability mass function is discrete," and thought, for example, that random and variable were redundant, that (0, 1) referred to a point in the plane, or that range means the analytic range from the concepts of domain, codomain, and range. Understanding statistics requires a strong foundation in analysis, algebra, probability theory, and set theory since it is an interdisciplinary off-shoot of mathematics. Indeed, a working knowledge of programming is helpful to apply the theory of statistics.

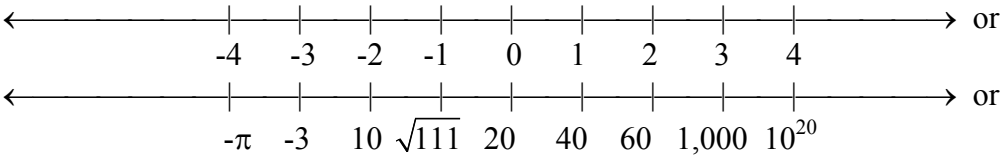
However, as is the case with most mathematics, a cursory understanding of the subject allows many to delve into the subject. One can then use technology to derive solutions to problems that might not ordinarily be easily obtained. Then inferences can be derived from the solutions which however pleasing or intuitively appealing are *dead wrong* and can cause mistakes, errors, confusion, etc.

It is the academic's responsibility to not fall into said trap. It is his responsibility to study an area of academia unencumbered with pre-concepts, notions, or biases. For if the search for knowledge is not entered into through the spirit of truth, honesty, honour, and curiosity, then how can one ever obtain knowledge, understanding, or - - one hopes - -wisdom?

So, we enter into an area free of pre-concepts, but demand that the axioms system we develop or study be consistent. One such example of a fundamental axiom system is the axioms of the real numbers. Recall from chapter two that the reals, $\mathbb{R}$, are the set of all points on the line and have graphical representation:



Recall there is ***no*** centre (e.g.: the nonsense about $\infty + (-\infty) = 0$) so the line is:



The axioms define the relationship between the points and define the binary operations addition and multiplication on the reals. They allow us to deduce many advanced properties of the reals

that are not necessarily true but are only conditionally true dependent on the truth of the axioms. So, in essence, the axioms of the reals are exemplary of that people which once thought were self-evident truths. However, there is no intrinsic truth to the axioms; they are but axioms that are generally considered of use and are generally agreed to.

Axiom Set 3.2.1: Let $U = \mathbb{R}$.. The **field axioms**[3] are:

Axiom 1 (closure of addition): $\forall\ x, y \in \mathbb{R}, x + y \in \mathbb{R}$ and $(x = w \wedge y = v) \Rightarrow (x + y = w + v)$.

Axiom 2 (commutative of addition): $\forall\ x, y \in \mathbb{R}, x + y = y + x$..

Axiom 3 (associative of addition): $\forall\ x, y, z \in \mathbb{R}, (x + y) + z = x + (y + z)$.

Axiom 4 (existence of identity of addition): $\exists$ a unique number $0\ \ni x + 0 = x \quad \forall\ x \in \mathbb{R}$.

Axiom 5 (existence of additive inverse): $\forall\ x \in \mathbb{R}\ \exists$ a unique number $-x\ \ni x + (-x) = 0$.

Axiom 6 (closure of multiplication): $\forall\ x, y \in \mathbb{R}, x \cdot y \in \mathbb{R}$ and $(x = w \wedge y = v) \Rightarrow (x \cdot y = w \cdot v)$.

Axiom 7 (commutative of multiplication): $\forall\ x, y \in \mathbb{R}, x \cdot y = y \cdot x$.

Axiom 8 (associative of multiplication): $\forall\ x, y, z \in \mathbb{R}, (x \cdot y) \cdot z = x \cdot (y \cdot z)$.

Axiom 9 (existence of identity of multiplication): $\exists$ a unique number $1\ \ni x \cdot 1 = x \quad \forall\ x \in \mathbb{R}$

where $1 \neq 0$.

Axiom 10 (existence of multiplicative inverse): $\forall\ x \in \mathbb{R}\ \ni x \neq 0\ \exists$ a unique number x^{-1}

$\ni x \cdot (x^{-1}) = 1$.

Axiom 11 (distributive of multiplication over addition): $\forall\ x, y, z \in \mathbb{R}, x \cdot (y + z) = (x \cdot y) + (x \cdot z)$.

Algebraists (of whom I am not one) study sets that have these properties and not that if the set along with two binary operators satisfy these axioms for that set, then the set is said to be a field (hence the terminology field axioms of the reals). Such systems will be studied in Math 371 and 372 (Abstract Algebra I and II). For now, it suffices to note that the student should have been exposed to these axioms prior to entry in college and convinced himself of their apparent veracity.

[3] These axioms form the foundation of Math 361 and 362 (Real Analysis I and II) along with the order axioms and the completeness axiom; all the interesting results from algebra, graph theory, functional analysis, differential, and integral Calculus all stem from these). When one studies other systems which have these properties they are studying Algebra (Math 371 and 372).

Axiom Set 3.2.2: Let $U = \mathbb{R}$. The **order axioms** are:

Axiom 12 (trichotomy): $\forall$ x, y $\in \mathbb{R}$, exactly one of the following relationships exists between x and y :

$x < y$, $x = y$, $\vee$ $x > y$. [meaning that $(x < y)$ exor $(x = y)$ exor $(x > y)$].

Axiom 13 (transitivity of "<"): $\forall$ x, y, z $\in \mathbb{R}$, $[(x < y) \wedge (y < z)] \Rightarrow (x < z)$.

Axiom 14 (preservation of order under addition): $\forall$ x, y, z $\in \mathbb{R}$, $(x < y) \Rightarrow (x + z < y + z)$.

Axiom 15 (preservation of order for positive multiplier): $\forall$ x, y $\in \mathbb{R}$, $[(x < y) \wedge (0 < z)] \Rightarrow$

$(x \cdot z < y \cdot z)$.

The order axioms establish the relationship between points and axiomatised the important principle that two distinct points occupy different geometric locations (the trichotomy), that order is transitive, that order is preserved under addition, and that order is preserved for a positive multiplier. Note that the theorem that order is reversed for a negative multiplier need not be assumed - - it can be proven based on these axioms!

Furthermore, that $0 \neq 1$ is all we assumed from the field axioms. So, we can prove certain established truths about some of the relationships with numbers that most students believe are true but have not proved are true. For example the fact that that $0 \cdot 0 = 0$, $0 \cdot x = 0$ where x is any real, $0 < 1$, and other titbits based on the field axioms and the order axioms.

Claim 3.2.1: $0 \cdot 0 = 0$.

Proof: Let $U = \mathbb{R}$. Assume the field axioms and order axioms of the reals. Consider that 0 is a unique real number by the axiom of additive identity. Let x be a real number. Note that $x + 0 = x$ by the axiom of additive identity. But, $x + 0 = 0 + x$ by the axiom of additive commutativity. So,
$0 + x = x$ by transitivity of equality. Now, $0 \cdot (0 + x) = 0 \cdot x$ by preservation of equality (closure) under multiplication. So, $0 \cdot 0 + 0 \cdot x = 0 \cdot x$ by the distributive axiom of multiplication over addition. Now, note $0 \cdot 0 = b$ for some real number b since $\mathbb{R}$ is closed under multiplication[4].

Also, since $\mathbb{R}$ is closed under multiplication this means that $0 \cdot x = a$ for some real number a.

So, we have $0 \cdot 0 + a = a$ by substitution.[5] Since $a \in \mathbb{R}$, -a exist and is real by the axiom of the additive inverse. So, we know that $0 \cdot 0 + a + (-a) = a + (-a)$ by the preservation of equality under addition. However, that means that $0 \cdot 0 + (a + (-a)) = a + (-a)$ by the associativity of

[4] We will use this later.
[5] This makes it easier to see the next part of the proof to substitute for $0 \cdot x$ but it is not necessary just as it was not necessary to substitute for $0 \cdot 0$.

addition. Note that a + (-a) = 0 by the axiom of the additive inverse. So, we have $0 \cdot 0 + (a + (-a)) = 0 \Rightarrow$
$0 \cdot 0 + 0 = 0$. Recall that $0 \cdot 0$ was b, so we have $b + 0 = 0$ by substitution. But the axiom of additive identity requires that $b + 0 = b$. So, by the transitivity of equality note that $b = 0$. Hence, $0 \cdot 0 = 0$.

QED

I am certain that there is a shorter more elegant and pleasing argument that establishes the veracity of $0 \cdot 0 = 0$. However, as stated in the previous chapters the object of mathematics at this level (or any level if one inquires as to my opinion) is not elegance by correctness. Try to create arguments which are true and blast the elegance!

Since the claim is true, let us state it as a lemma (a small theorem, recall, that helps prove a 'bigger' result subsequently):

Lemma 3.2.1: $0 \cdot 0 = 0$.

Let us consider another ditty that many people believe is a definition but is not.

Claim 3.2.2: $0 \cdot x = 0 \; \forall \; x \in \mathbb{R}$.

Proof: Let $U = \mathbb{R}$. Assume the field axioms and order axioms of the reals. By the existence of the additive identity axiom, 0 exists. Note that -0 is 0 since 0 has an additive inverse. We have already established $0 \cdot 0 = 0$ so, let x be a real number (meaning $x \in \mathbb{R}$) where $x \neq 0$. Note that $0 \cdot x$ is a real number since the reals are closed under multiplication (meaning it exists) so $\exists$ c $\in \mathbb{R}$

$\ni 0 \cdot x = c$. Note that (-c) is also a real number by the axiom of the additive inverse.
Consider that $0 \cdot x = (0 + 0) \cdot x$. Now, $0 \cdot x = (0 \cdot x) + (0 \cdot x)$ by the distributive axiom of multiplication over addition (and commutativity since the axiom was stated for multiplication on the left). But $0 + 0 \cdot x = 0 \cdot x$ by the axiom of the additive identity. So, we have
$0 + 0 \cdot x = (0 \cdot x) + (0 \cdot x)$. Thus, $0 + c = c + c$. Now we can say that this implies that
$0 + c + (-c) = c + c + (-c)$. But addition is associative so, $0 + c + (-c) = c + c + (-c) \Rightarrow$
$0 + (c + (-c)) = c + c + (-c) \Rightarrow 0 + (c + (-c)) = c + (c + (-c)) \Rightarrow 0 + 0 = c + 0 \Rightarrow 0 = c$.
Since $c = 0 \cdot x$ we finally note that it must be the case that $0 = 0 \cdot x \; \Rightarrow \; 0 \cdot x = 0$.

QED

Since the claim is true, let us state it as a theorem:

Theorem 3.2.1: $0 \cdot x = 0 \; \forall \; x \in \mathbb{R}$.

Claim 3.2.3: $(-1) \cdot x = -x \; \forall \; x \in \mathbb{R}$.

Proof: Let $U = \mathbb{R}$. Assume the field axioms and order axioms of the reals. By the existence of the multiplicative identity axiom, 1 exists. By the additive inverse axiom, (-1) exists. Now let x be a real number. By the additive inverse axiom, (-x) exists.

Consider that the reals are closed under multiplication, so $\exists\, d \in \mathbb{R} \ni (-1) \cdot x = d$. Now we know that d = d; so, $(-1) \cdot x = (-1) \cdot x$. Consider $x + (-1) \cdot x = x + (-1) \cdot x$ because equality is preserved under addition. But, $x = 1 \cdot x$.
So, $1 \cdot x + (-1) \cdot x = x + (-1) \cdot x$. Since multiplication over addition is distributive, we have $(1 + (-1)) \cdot x = x + (-1) \cdot x$. But 1 + (-1) = 0; so we have $0 \cdot x = x + (-1) \cdot x$.. However, by theorem 3.2.1, we therefore have $0 = x + (-1) \cdot x$.. Now, the axiom of the additive inverse states that the additive inverse is unique; so since –x is the additive inverse of x it therefore follows
that $-x = + (-1) \cdot x$..

QED

Since the claim is true, let us state it as a theorem:

Theorem 3.2.2: $(-1) \cdot x = -x \;\forall\, x \in \mathbb{R}$.

Claim 3.2.4: $0 < 1$.

Proof: Let $U = \mathbb{R}$. By the existence of the additive identity axiom, 0 exists. By the existence of the multiplicative identity 1 exists. Also, $0 \neq 1$ by the same axiom. Now by the trichotomy axiom exactly one of the following relationships exists $0 < 1$ or $0 = 1$, or $0 > 1$.
Case 1: $0 = 1$. But $0 \neq 1$. Hence, we have a contradiction. So, this cannot be the case.
Case 2: $0 > 1$. Now, 1 has an additive inverse, (-1), which is unique by the axiom of additive inverses. So, 1 + (-1) = (-1) + 1 = 0 since addition is commutative.
Consider $0 > 1 \Rightarrow 0 + (-1) > 1 + (-1)$ by the preservation of order under addition. Nonetheless, 0 + (-1) = (-1) + 0 since addition is commutative, and by the existence of additive identity axiom, we therefore know that 0 + (-1) = -1. So, $0 > 1 \Rightarrow 0 + (-1) > 1 + (-1) \Rightarrow (-1) + 0 > 1 + (-1) \Rightarrow$
$0 + (-1) > 1 + (-1) \Rightarrow (-1) > 1 + (-1) \Rightarrow (-1) > 0$. So, (-1) is a positive number; meaning that $0 < (-1)$. Now, let us consider that $1 < 0$ and $0 < (-1)$. Applying the axiom of preservation of order for a positive multiplier where x = 1, y = 0, z = -1 yields $1 \cdot (-1) < 0 \cdot 0$. Since multiplication is commutative, we have $(-1) \cdot 1 < 0 \cdot 0$. By lemma 3.2.1 we therefore know that $(-1) \cdot 1 < 0$.
But by the axiom of the multiplicative identity $(-1) \cdot 1 = -1$. So, $-1 < 0$. However, we also have $0 < -1$. So, $-1 < 0 \wedge 0 < -1$ which is a contradiction of the trichotomy axiom; so, this cannot be the case.
So, we *must* conclude $0 < 1$.

QED

Since the claim is true, let us state it as a corollary (a small theorem, recall, that follows from some 'bigger' result):

Corollary 3.2.1: $0 < 1$.

There is a host of other results that one might view as 'trivial' but are in fact important results of field and order axioms of the reals such that when we were young we considered them 'given' but are in fact results that are proven based on the axioms. As such they need not be assumed (which is not to say that many don't assume them[6]) they are deduced from the axioms. So, they rely on the axioms but are different than the axioms insofar as the axioms are those statements assumed to be the case that are not deduced from other statements.

Now, there is yet another axiom of the reals that is of import. That is the completeness axiom. It is very useful in analysis. We shall not define the terms in this axiom; we shall simply note that it is an axiom that will be used and results of which shall be studied in Math 361 (Real Analysis I). If you have an inkling as to the definition of boundedness and supremum, fine, that is not our concern at this stage of our mathematical development.

Axiom 16 (completeness): $\forall A \subseteq \mathbb{R} \ni A$ is bounded above $\exists$ a number m which is the supremum of the set.

Recall in chapter two we discussed some basic notions of sets, properties of sets, Venn diagrammes, etc. Our treatment of sets was (to say the least) elementary. However, the subject was developed with a naïveté that was purposeful. In the theory of methods of teaching (the edu-speak term is pedagogy), one perspective is called constructivism. That is to say that the student should construct the ideas through his work. I do not adhere to said method, but liberally borrow from it from time to time when prudence and experience dictate. To study in an expository manner sets axiomatically is not the best way (in my opinion, though I lean somewhat toward the axiomatic side of methods of teaching arguments); it seems most students learn the theory of sets better by first getting some applied experience then studying the subject in a deep vertical manner later.

Nonetheless, the axioms exist. To leave you hanging in thin air to 'develop' them on your own might be an interesting exercise; but would be time-consuming and might not yield the desired result of the student learning to apply set theoretic concepts to algebra, analysis, probability, etc. Hence, we shall note the axioms, discuss a couple of the axioms, and leave it at that.

[6] Read this carefully. Decide what is being said.

Axiom Set 3.2.3: Let *U* be a well defined universe. Let A, B, and C be sets of elements from that universe. The **axioms of set theory**[7] are:

Axiom 1 (The Axiom of Extension) Two sets are equal iff they have the same elements.

Axiom 2 (The Axiom of Null) There exists a set with no elements, call it $\varnothing$.

Axiom 3 (The Axiom of Pairing) Given any sets A and B, there exists a set C whose elements are A and B.

Axiom 4 (The Axiom of Union) Given any set A, the union of all elements in A is a set.

Axiom 5 (The Axiom of Power Set) Given any set A, there exists a set B consisting of all the subsets of A.

Axiom 6 (The Axiom of Separation) Given any set A and a sentence p(a) that is a statement for all $a \in A$, then there exists a set $B = \{a \in A: p(a) \text{ is true}\}$.

Axiom 7 (The Axiom of Replacement) Given any set A and a function *f* defined on A, the image *f*(A) is a set.

Axiom 8 (The Axiom of Infinity) There exists a set A such that $\varnothing \in A$, and whenever $a \in A$, it follows that $a \cup \{a\} \in A$.

Axiom 9 (The Axiom of Regularity) Given any non-empty set A, there exists an $a \in A$ such that $a \cap A = \varnothing$.

Axiom 10 (The Axiom of Choice) Given any non-empty set A whose members are pair-wise disjoint non-empty sets, there exists a set B consisting of exactly one element taken from each set belonging to A.

Now, notice the axiom of extension is a convenient way to describe a set so that the concept of $U = \mathbb{N}$ (or any set but let us use this example) means that if one were to say that A = {1, 2, 3} was a set and B = {2, 1, 3} was a set; it can be determined by the axiom of extension that naming them two different things does not imply that they are different!

Note the axiom of null; its existence is axiomatic. One can study[8] other systems of set theory that, for example, do not admit the axiom of null. So, one can readily opine that $A \cap A^C$ presents a problem. In actuality it is not so difficult for $A \cap A^C$ to not exist in such a system.

[7] These are the axioms that are the basis of Math 255, Set Theory. However, they are difficult to 'deal' with and so one does not typically work with them until a graduate class in Set Theory.

[8] I have studied such set theories. It is a fascinating discussion to determine when one deletes certain axioms what results hold or do not hold. That is to say that just because an axiom is presented does not necessitate its acceptance. So, one can study mathematical systems such that different axiom systems are predicated. However, if such a system is predicated the axioms must be stated.

The other axioms are decidedly more complex and sublime so we shall leave the discussion of those for a later course. Let us instead consider another axiom system, the axioms of probability.

Axiom Set 3.2.4: Let U be a well defined universe from set theory. Rename the universe S and call it a sample space. Let E, E_i, F, etc. be sets of elements from that sample space such that E, E_i, F, etc. will be called events. The **axioms of probability**[9] are:

Let S denote the sample space, E, E_i, F, etc. events and the notation Pr(•) the probability of whatever [the dot is a dummy for an event].

Axiom 1 S is the space $\Rightarrow \Pr(S) = 1$.

Axiom 2 E is an event $\Rightarrow 0 \le \Pr(E) \le 1$.

Axiom 3 Let I be an index set. The collection $\{E_i\}_{i \in I}$ being mutually exclusive

$$\Rightarrow \Pr(\bigcup_{i \in I} E_i) = \sum_{i \in I} \Pr(E_i)$$

Note that there are definitions that should be clearly delineated before the discussion of the axioms. For example, we did define sample space and event, but not mutual exclusivity. Two events are mutually exclusive iff their intersection is null. One must also define summation, function, etc. for compleat understanding of these axioms. One should clearly realise that the axioms of probability rest heavily on the reader's understand of sets. I dare say that most upper division course-work relies on a clear, unambiguous, and full understanding of logic and sets. They are the tools that assist the student in mastering upper-level mathematics.

We shall return to these axioms later in the course when we discuss introductory probability and statistics. We will use them and in so doing (hopefully) a better understanding will develop as to their importance and utility.

[9] These are the axioms that form the basis of Math 341 and Math 342 (Probability and Mathematical Statistics). Note: the axioms of probability are one of the shortest lists I can recall for an area of mathematics - - they form the basis of probability theory; however, probability theory depends on other mathematical theories (for example, set theory). Thus, the brevity of the axiom list is *somewhat* deceiving.

§ 3.3 A BIT OF FORMAL NATURAL ARITHMETIC.

Imagine that we are children and we are learning about mathematics. Would we first discuss calculus? Of course not; that would be preposterous. We would begin with the natural numbers and arithmetic. We would conceive of 1, 2, 3, and so forth. We would then begin to understand things like 1 and 1 is 2. Later we would be introduced to 1 + 1 = 2 and be instructed that '+' signifies the concept of 'and' (probably thinking of one and one apple makes two apples, I suppose) whilst the '=' means the verb 'is.' We would memorise addition and multiplication tables, then develop subtraction, division, the rational numbers, etc. That is how our 'numerical sense' was probably developed, refined, and enhanced. But we are *not* children, so let us discuss the natural numbers in a more rigorous way.

Axiom Set 3.3.1: Let $U = \mathbb{N}$.

Axiom 1 (closure of addition): $\forall\, x, y \in \mathbb{N}, x + y \in \mathbb{N}$.

Axiom 2 (commutative of addition): $\forall\, x, y \in \mathbb{N}, x + y = y + x$.

Axiom 3 (associative of addition): $\forall\, x, y, z \in \mathbb{N}, (x + y) + z = x + (y + z)$.

Axiom 4 (axiom of identity): $\forall\, x \in \mathbb{N}, x = x$.

Axiom 5 (axiom of substitution): $\forall\, x, y \in \mathbb{N}, x = y \Rightarrow y = x$.

Axiom 6 (axiom of transitivity of equality): $\forall\, x, y, z \in \mathbb{N}, (x = y \wedge y = z) \Rightarrow x = z$.

Axiom 7 (preservation of equality): $\forall\, w, x, y, z \in \mathbb{N}, (x = w \wedge y = z) \Rightarrow (x + y = w + z)$.

One can prove simple theorems based on these axioms. For example consider the following theorem:

Theorem 3.3.1: If x, y, and z are natural numbers, then (x + y) + z = (z + x) + y.

Proof: Let $U = \mathbb{N}$. Assume axiom set 3.3.1. Let x, y, and z be natural numbers.

Consider (x + y) + z. By the axiom of closure of addition $x + y \in \mathbb{N}$. So, it is some natural number. Hence that natural number plus z is also a natural number by the closure of addition axiom. Now, we know that (x + y) + z = x + (y + z) by the axiom of associativity.
By the axiom of commutativity, x + (y + z) = x + (z + y).
By the transitivity of equality we know that (x + y) + z = x + (z + y).
By the axiom of associativity we know x + (z + y) = (x + z) + y.
By the transitivity of equality we know that (x + y) + z = (x + z) + y.
By the axiom of commutativity, (x + z) + y = (z + x) + y.
By the transitivity of equality we know that (x + y) + z = (z + x) + y.
So, we *must* conclude (x + y) + z = (z + x) + y.

QED

Other trivial theorems can be deduced from the axioms. It is important to realise however, that the reasoning used is sound and that when we do something as trite as the following:

$$\begin{array}{r} 7 \\ 5 \\ +\ 6 \\ \hline 18 \end{array}$$

we are implicitly using the associative and commutative axioms!

Note the first three basic axioms of $(\mathbb{N}, +)$ [just a fancy way of saying the natural numbers with addition). Do they hold for subtraction? In other words consider the following 'axioms;' are they axioms (e.g.: must we assume them) and are they even true?

'Axiom' 1 (closure of subtraction): $\forall\ x, y \in \mathbb{N}, x - y \in \mathbb{N}$.

'Axiom' 2 (commutative of subtraction): $\forall\ x, y \in \mathbb{N}, x - y = y - x$.

'Axiom' 3 (associative of subtraction): $\forall\ x, y, z \in \mathbb{N}, (x - y) - z = x - (y - z)$.

Note that axiom 4, 5, and 6 do not explicitly mention addition. Hence, do they hold for subtraction? Why or why not? Finally, what of axiom 7?

The section on formal arithmetic will be expanded in subsequent semestres. However, for now it should be clear as to the nature of our discussion. Much is assumed but does it really have to be assumed or is it that we have not elucidated the reasons behind our actions?

§ 3.3 EXERCISES.

1. Name the axiom or axioms that justify the following.

A. $7 + 11 = 11 + 7$.
B. $6 + (7 + 5) = (6 + 7) + 5$.
C. $6 + (7 + 5) = 7 + (6 + 5)$.
D. Let x and y be natural numbers. $x + y = x + y$

2. Sketch an argument (some would call this an informal proof) to argue the veracity of the following based on axiom set 3.3.1:

A. If w, x, y, and z are natural numbers, then $[(x + y) + z] + w$ is a natural number.
B. If w, x, y, and z are natural numbers, then $x + y + z + w$ is a natural number.
C. If w, x, y, and z are natural numbers, then $(x + y) + (z + w) = (z + w) + (x + y)$.
D. If x, y, and z are natural numbers, then $(x + y) + z = z + (y + x)$.
E. If w, x, y, and z are natural numbers, then $[(x + y) + z] + w = (x + y) + (z + w)$.
F. If w, x, y, and z are natural numbers, then $[(x + y) + z] + w = w + ([y + x] + z)$.

3. Note the seventh basic axioms of $(\mathbb{N}, +)$. Does it hold for subtraction? Why or why not?

§ 3.4 ANOTHER TYPE OF ARITHMETIC.

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 ($\mathbb{N}^* = \{0, 1, 2, 3, 4, 5, \ldots\}$) 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 rigor 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 7 in base 5? Thus 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 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.
Now, 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 $(1 \times 2^4) + (1 \times 2^3) + (0 \times 2^2) + (1 \times 2^1) + (0 \times 2^0)$.
Hopefully, it is facile to see that we therefore have (16) + (8) + (0) + (2) + (0).

So, $11010_2 \equiv 26$ 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.

So, suppose we are presented with 113241_5. What is it?

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

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 and multiplication it is logical to conclude that this process will require division (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.

$$\begin{array}{rrl}
 & 1296\overline{)3401}\;\;2 & \text{Note that } 1{,}296 \times 2 = 2{,}596. \text{ So, we subtract} \\
 & \underline{-2596} & \\
 & 805 & \text{leaving 805 as a remainder.} \\
\text{Now,} & 216\overline{)805}\;\;3 & \text{Note that } 216 \times 3 = 648. \text{ So, we subtract} \\
 & \underline{-\;648} & \\
 & 157 & \text{leaving 157 as a remainder.} \\
\text{Now,} & 36\overline{)157}\;\;4 & \text{Note that } 36 \times 4 = 144. \text{ So, we subtract} \\
 & \underline{-\;144} & \\
 & 13 & \text{leaving 13 as a remainder.} \\
\text{So,} & 6\overline{)13}\;\;2 & \text{Note that } 6 \times 2 = 12. \text{ So, we subtract} \\
 & \underline{-\;12} & \\
 & 1 & \text{leaving 1 as a remainder.} \\
\text{Finally,} & 1\overline{)1}\;\;1 & \text{Note that } 1 \times 1 = 11. \text{ So, we subtract} \\
 & \underline{-\;1} & \\
 & 0 & \text{leaving no remainder.}
\end{array}$$

Hence, it is the case that $3401 = 2596 + 648 + 144 + 12 + 1 =$
$(2 \times 1296) + (3 \times 216) + (4 \times 36) + (2 \times 6) + (1 \times 1) =$
$(2 \times 6^4) + (3 \times 6^3) + (4 \times 6^2) + (2 \times 6^1) + (1 \times 6^0) \Rightarrow$
$3401 \equiv 23421_6$.

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

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. We only need begin our work with 512. 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

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.

$144\overline{)711}$ Note that $144 \times 4 = 576$. So, we subtract

-576

135 leaving 135 as a remainder.

Now, $12\overline{)135}$ Note that $12 \times 11 = 132$. But, we cannot use '11' so we use E.

-132 Subtract

3 leaving 3 as a remainder.

Now, $1\overline{)3}$ Note that $1 \times 3 = 3$ So, we are done.

Hence, $711 \equiv 4E3_{12}$.

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

For example converting 312_4 to base 7 would entail considering that 312_4 is $(3 \times 4^2) + (1 \times 4^1) + (2 \times 4^0) = (3 \times 16) + (1 \times 4) + (2 \times 1) = 48 + 4 + 2 = 54$. 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.

$49\overline{)54}$ Note that $49 \times 1 = 49$. So, we subtract

-49

5 leaving 5 as a remainder.

$7\overline{)5}$ Note that $7 \times 0 = 0$. So, we don't subtract.

still leaving 5 as a remainder.

Now, $1\overline{)5}$ Note that $1 \times 5 = 5$. Subtract

-5

0 So, we are done.

Hence, $312_4 \equiv 105_7$. Please note that we are using transitivity to deduce this.

There are shortcuts 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!

§ 3.4 EXERCISES.

1. Convert the following numbers to base ten.

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

2. Convert 913 the following bases

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

3. Convert the following numbers to the specified base.

A. 813_9 to base 3
B. 110311_4 to base 2
C. 110101_2 to base 3
D. 03413_5 to base 8
E. 110311_7 to base 12
F. 1661_8 to base 7
G. 313_4 to base 5
H. 3103013_4 to base 9
I. 1010101_2 to base 8
J. 34013_5 to base 7
K. 110131_7 to base 5
L. 6161_8 to base 4
M. 6611_8 to base 6
N. 1166_8 to base 6
O. 6116_8 to base 6
P. $3ET1_{12}$ to base 2
Q. 1562_{12} to base 5
R. 111_{11} to base 5

4. Convert 1100101010101010101_2 to base 10.

5. Let the digits for base 16 be {0, 1, 2, 4, 5, 6, 7, 8, 9, A, B, C, D, E, F} where A is ten B is eleven, C is twelve, D is thirteen, E is fourteen, and F is fifteen. Convert the following to base 10:

A. 813_{16}
B. 110311_{16}
C. BAD_{16}
D. $5F3_{16}$
E. $FACE_{16}$
F. 111_{16}

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

7. 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 $z \equiv d_b$
$a_b \times c_b = f_b$ if and only if $a_b \equiv x_{10}$ and $c_b \equiv y_{10}$ yields $x \times y = w$ and $z \equiv f_b$).
Compute the following:

A. $813_9 + 712_9$
B. $110311_4 + 110311_4$
C. $110101_2 + 110101_2$
D. $110101_2 + 1101_2$
E. $110101_2 + 1001_2$
F. $110101_2 + 1111_2$
G. $31_4 + 11_4$
H. $31_4 + 32_4$
I. $31_4 \times 11_4$
J. $31_4 \times 32_4$
M. $1101_2 \times 1001_2$
N. $6_8 \times 7_8$
O. $3_8 \times 2_8$
P. $3ET1_{12} \times 7_{12}$
Q. $110101_2 \times 1001_2$

§ 3.5 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, 366, or 367 days?[10]

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.

So, let us define modular arithmetic as follows. Let $U = \mathbb{N}^*_p$ and $D = \mathbb{N}^*_{(p-1)}$ Where $p \in \mathbb{N}$.

Let $a \in D$ and $b \in D$. Let $a +_p b$ be $z \in D$ where z is the remainder from the division of p into (a + b) under ordinary addition for the natural numbers.

So, for example let us see what $+_4$ is. $0 +_4 0 = 0$, $0 +_4 1 = 1$, $0 +_4 2 = 2$, $0 +_4 3 = 3$, $1 +_4 0 = 0$, $2 +_4 0 = 2$, $3 +_4 0 = 3$ all because 4 divided into any of these is zero with remainder whatever the non-zero constant was. Consider $1 +_4 1 = 2$, $2 +_4 1 = 3$, $1 +_4 2 = 3$ 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.

$3 +_4 2 = 1$ since $5 \div 4 = 1$ with remainder 1.

$3 +_4 3 = 2$ since $6 \div 4 = 1$ with remainder 2.

Note that this is all good and well but it can be confusing. So, let us note the results in tabular form:

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

This should look strikingly familiar (see section 3.1).

[10] Check this out – hint: years ending with '00' but not divisible by 400.

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

$+_3$	0	1	2
0	0	1	2
1	1	2	3
2	2	3	0

and:

$+_6$	0	1	2	3	4	5
0	0	1	2	3	4	5
1	1	2	3	4	5	0
2	2	3	4	5	0	1
3	3	4	5	0	1	2
4	4	5	0	1	2	3
5	5	0	1	2	3	4

We can also define modular multiplication: Let $U = \mathbb{N}_p^*$ for some $p \in \mathbb{N}$.

Let $a \in \mathbb{N}_p^*$ and $b \in \mathbb{N}_p^*$. Let $a \times_p b$ be $z \in \mathbb{N}_p^*$ where z is the remainder from the division of p into $(a \times b)$ under ordinary multiplication for the natural numbers. Do the arithmetic in your head for mod 4 multiplication to get the following:

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

Now reconsider 11 + 5 = 4. Recall we are *not* in the business to play 'games.'
Consider $+_{12}$. Note it is fine to consider $U = \mathbb{N}_{12}^*$ and let the set of digits be $\mathbb{N}_{11}^*$.
$11 +_{12} 5 = 4$ 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.). 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 there exist a distributive property of modular multiplication over modular addition? If so, how would you prove it; if not, can you devise a counterexample?

It is perfectly acceptable to define a system of modular arithmetic that includes $p \in \mathbb{N}$ for $U = \mathbb{N}_p^*$ as with time. Hence, in modular 12 arithmetic, $7 +_{12} 5 = 0$ but in practical applications most would say it is 12 Since $12 \Leftrightarrow 0$ 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.

§ 3.5 EXERCISES.

1. Construct the table for addition and multiplication mod 5.

2. Construct the table for addition and multiplication mod 8.

3. Calculate the following

A. $8 +_9 1 +_9 3$
B. $8 \times_9 1 \times_9 3$
C. $1 +_2 1 +_2 1$
D. $0 \times_2 1 \times_2 1$
E. $4 +_5 3 +_5 1$
F. $4 \times_5 3 \times_5 1$
G. $7 +_9 7 +_9 7$
H. $7 \times_9 7 \times_9 7$
I. $7 +_8 7 +_8 7$
J. $7 \times_8 7 \times_8 7$
K. $3 +_5 3 +_5 1$
L. $4 \times_5 3 \times_5 3$
N. $7 +_8 5 +_8 6$
O. $6 +_8 5 +_8 7$
N. $(7 +_8 5) +_8 6$
O. $5 \times_8 7 \times_8 6$
N. $5 +_{12} 5 +_{12} 5$
O. $5 \times_8 6 \times_8 7$

4. What is incorrect about the following claim?
$9 \times_8 7 = 7 \times_8 9$

5. Compute the following:

A. $813 +_9 712$
B. $110311 +_4 110311$
C. $110101 +_2 110101$
D. $110101 +_2 1101$
E. $110101 +_2 1001$
F. $110101 +_2 1111$
G. $31 +_4 11$
H. $31 +_4 32$
I. $31_4 \times_4 11$
J. $31 \times_4 32$
M. $1101 \times_2 1001$
N. $6 \times_8 7$
O. $3 \times_8 2$
P. $31 \times_{12} 7$
Q. $110101 \times_2 1001$

6. Compare the results of exercise 5 with exercise 7 of section 3.4. 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?

§ 3.6 THE PEANO AXIOMS, COUNTING, AND MATHEMATICAL INDUCTION.

We learned to count before elementary school (one hopes); but, the formal theory of counting is oft called Number Theory (Math 475) or Combinatorics (not yet offered).

At an introductory level for mathematics, combinatorics is considered a branch of Discrete Mathematics (Math 211) in which the main focus is the number of ways to choose or arrange objects from a finite set from Set Theory (Math 255). It is a branch of number theory insofar as the axioms of number theory create the building blocks from which combinatorics arises. Much work in Numerical Analysis (Math 467) requires a rudimentary understanding of number theory as well as Real Analysis (Math 361 – 362 – 463 sequence). Indeed one needs to understand theoretical mathematics (including counting theory) in order to really understand and master Applied Mathematics) Math 325 and 327).

Most of the counting techniques that we will concern ourselves with are of the type that lay the groundwork for an intuitive understanding of Probability Theory (Math 341 - 342).

We have already discussed at least one method of counting from a finite set and that was the applications of Venn Diagrammes to surveys. Now, we will extend our understanding to some more complex problems.

Recall that $\mathbb{N} = \{1, 2, 3, 4, 5, \ldots, (k - 1), k, (k + 1), \ldots.\}$. They seemingly go on and on – well, they do because it has been *proven* to be true. Let us consider this important theorem.

Theorem 3.6.1 (the Archimedean property of $\mathbb{N}$ in $\mathbb{R}$): The natural numbers are unbounded above in the reals.

The properties of addition of natural numbers can be derived from a short set of axioms. The axioms are called **the Peano Axioms**:

There exists a set, P, which is defined by the following four axioms.

Axiom 1: There exists a natural number, call it 1, that is not the successor of any other natural number.

Axiom 2: Every natural number has a unique successor. If $k \in P$, then let k' denote the successor of k.

Axiom 3: Every natural number except one is the successor of exactly one natural number.

Axiom 4: If M is **a** set of natural numbers such that
(i) $1 \in M$ and
(ii) for each $k \in P$, if $k \in M$, then $k' \in P$,
then $P = M$.

P, of course is $\mathbb{N}$.

So, the Peano axioms assert the uniqueness of the naturals that this successor property along with the element 1 creates the entirety of the natural numbers. No matter how you name the set (you can call it Ray, or you can call it Jay, . . .) if it has these properties then it really is the naturals.

From these axioms arise the natural numbers by defining what addition by one means.

Definition 3.6.1: For every $k \in \mathbb{N}$, define $k + 1 = k'$.

Then, note inductively, the entire understanding of addition flows from this definition (likewise multiplication, etc.).

So, it seems the basis of our understanding of counting will be based $\mathbb{N}$.

No, pity. We need $\mathbb{N}^*$.

Recall $\mathbb{N}^* = \{ 0, 1, 2, 3, \ldots, (j - 1), j, (j + 1), \ldots . \} = \mathbb{N} \cup \{0\} = \{0\} \cup \mathbb{N}$.

Finally, before proceeding we need to define factoral.
Let $k \in \mathbb{N}^*$ recursively define k! Such that

$0 ! = 1$
$1 ! = 1$
$2 ! = 2 \cdot 1$
$3 ! = 3 \cdot 2 \cdot 1$
.
.
.
$k ! = k \cdot (k - 1) \cdot \ldots \cdot 3 \cdot 2 \cdot 1$ where $k \geq 3$.

A more succinct definition is $k ! = \prod_{j=1}^{k} j$

Theorem 3.6.2: Let $k \in \mathbb{N}$. It is the case that $k ! = k \cdot (k - 1) !$

Now, to the business at hand:

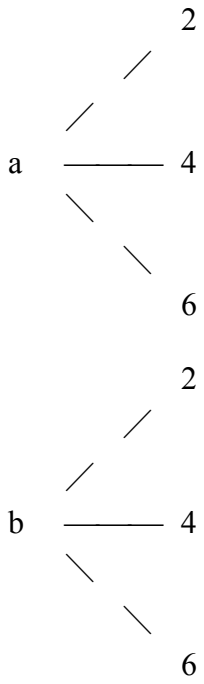
Lemma 3.6.3: If we have a set, A, with m elements such that $m \in \mathbb{N}$ and we have a set, B, with n elements such that $n \in \mathbb{N}$, then the number of ways to order the elements from A then the elements from B is $m \cdot n$.

Theorem 3.6.3 (The Fundamental Principle of Counting): If activities 1, 2, 3, . . ., k can be performed in $n_1, n_2, n_3, \ldots n_k$ ways, respectively, such that $k \in \mathbb{N}$ and $n_i \in \mathbb{N} \; \forall \; i \in \mathbb{N}_k$ then the k activities can be performed in $\prod_{j=1}^{k} n_j = n_1 \cdot n_2 \cdot n_3 \cdot \cdots \cdot n_k$ ways.

Consider we choose one of 2 objects from the set {a, b}, then we choose one of three objects from the set {2, 4, 6}. Hence, the number of ways to do this is $2 \cdot 3 = 6$.

The sequence of activities can be illustrated with a set of ordered pairs since the activities are in order:
{(a, 2), (a, 4), (a, 6), (b, 2), (b, 4), (b, 6)}.

Further, the sequence of activities can be illustrated with a **tree diagramme**:



A tree diagramme is a simple graphical illustration of an ordered sequence of activities.

In many counting problems, the task assigned is one involving arranging a set of objects. Now, the arrangement may or may not involve order.

Definition 3.6.2: Suppose the set A has n objects such that $n \in \mathbb{N}^*$ and we wish to order k of the objects $\ni k \leq n$ where $k \in \mathbb{N}^*$
The number of ways to do this is referred to as the **permutations of n things taken k at a time** and is symbolised as $P_{n,k}$ where $P_{n,k} = \frac{n!}{(n-k)!}$

Alternate notation for permutations include the following: $P_{n,k} = {}_nP_k = P_k^n = {}_k P^n = P(n, k)$.

Definition 3.6.3: Suppose the set A has n objects such that $n \in \mathbb{N}^*$ and we wish to choose k of the objects (without regard to order) $\ni k \leq n$ where $k \in \mathbb{N}^*$ The number of ways to do this is

referred to as the **combinations of n things taken k at a time** and is symbolised as $\binom{n}{k}$ where $\binom{n}{k} = \frac{n!}{(n-k)!\,k!}$

Alternate notation for combinations include the following: $\binom{n}{k} = C_{n,k} = {}_nC_k = C^n_k = {}_kC^n = C(n, k)$.

Theorem 3.6.4: Let $n \in \mathbb{N}^*$ and $k \in \mathbb{N}^* \ni k \leq n$. $P_{n,k} = k! \cdot \binom{n}{k}$.

As we know, inductive reasoning is fraught with fallacies. *Deductive reasoning is not.* Consider a statement $\Phi(n)$ that we wish to prove for all $n \in \mathbb{N}$. This is logically equivalent to the statement $\forall\, n \in \mathbb{N},\ \Phi(n)$. Such types of claims oft 'pop up' in number theory, analysis, algebra, etc. Thus, it would behoove us to have a method of proof to tackle such quantification claims. Well, fortunately there is a method:

First principle (axiom) of mathematical induction:
$\{ \Phi(1) \wedge [\forall\, k \in \mathbb{N}, \Phi(k) \rightarrow \Phi(k + 1)] \} \Rightarrow \forall\, n \in \mathbb{N},\ \Phi(n)$.

Thus, if we can prove the antecedent of the axiom, then by the axiom (an application of modus ponens) , we can deduce $\forall\, n \in \mathbb{N},\ \Phi(n)$.

Thus, there are two steps to prove a claim:

Prove $\Phi(1)$ is true. This is oft referred to as the basis step.
Prove given $k \in \mathbb{N}$, $\Phi(k) \rightarrow \Phi(k + 1)$. This is oft referred to as the inductive step.

To understand this process intuitively, note we have deduced an endless sequence of sentences.

$\Phi(1)$.
$\Phi(1) \rightarrow \Phi(2)$.
$\Phi(2) \rightarrow \Phi(3)$.
$\Phi(3) \rightarrow \Phi(4)$.
.
.
.
$\Phi(j - 1) \rightarrow \Phi(j)$.
$\Phi(j) \rightarrow \Phi(j + 1)$.
.
.
.

So, what we have is an endless sequence of modus ponens arguments:

$\Phi(1)$.		$\Phi(2)$.		$\Phi(3)$.	. . .	$\Phi(j)$.	. . .
$\underline{\Phi(1) \rightarrow \Phi(2).}$		$\underline{\Phi(2) \rightarrow \Phi(3).}$		$\underline{\Phi(3) \rightarrow \Phi(4).}$		$\underline{\Phi(j) \rightarrow \Phi(j + 1).}$	
$\therefore \Phi(2).$	$\therefore$	$\Phi(3).$	$\therefore$	$\Phi(4).$	. . . $\therefore$	$\Phi(j + 1)$. . . .	

Producing the endless sequence of true statements: $\Phi(1), \Phi(2), \Phi(3), \Phi(4), \ldots, \Phi(j), \ldots$

Note there is no need to begin with n = 1. What is necessary is that the set which we are considering is of the form of a "list" which in Set Theory will be defined for you and whose properties will be explored (such sets are called denumerable). The set $\mathbb{N}$ in the first principle of mathematical induction is referred to as the index set or the set of indices.

Also, note that since there is not a need to begin with n = 1, there will be other principles of mathematical induction. All generally conform to this idea; however, so understanding this basic form is a step toward understanding more sophisticate forms.

<u>Second principle (axiom) of mathematical induction</u>: Let $a_i \in I$, where I is some denumerable index set. $\{ \Phi(a_1) \wedge [\forall k \in \mathbb{N}, \Phi(a_k) \rightarrow \Phi(a_{(k+1)}] \} \Rightarrow \forall a_n \in I, \Phi(a_n)$.

<u>Third principle (axiom) of mathematical induction (*also called strong induction*)</u>: Let X be a subset of the natural numbers. If $k \in X$ when $\forall j \in \mathbb{N}$ ∋ $j < k$ $j \in X$, then $X = \mathbb{N}$.

Now, let us consider a claim.

Claim 3.6.1: $\sum_{j=1}^{n} j\cdot(j!) \le (n+1)!$ for all $n \in \mathbb{N}$. Prove or disprove the claim

As always, you must first read the claim and decide whether or not you think it is true (you may be wrong, but you have to practice this step; it is based on your prior experience and knowledge). It is an inductive step; hence, there is no guarantee that you are right.

Next, after considering the claim, suppose we think it true. *Thinking* it is true is not *proving* it is true. Hence, we need to construct a proof. We must announce it is a proof and frame it at the
beginning (Proof:) and at the end (Q.E.D.[Quod Erat Demonstratum]).

Proof:

0. Assume the premises (the axioms)	0. premises
1. Let n = 1	1. Basis step
2. Consider $\sum_{j=1}^{1} j(j!)$	2. Hypothesis
3. = 1 (1!)	3. Definition of Sigma
4. = 1 · 1	4. Definition of factoral
5. = 1	5. Multiplication
6. Consider (n + 1)!	6. Hypothesis
7. = (1 + 1)!	7. Substitution
8. = 2!	8. Addition
9. = 2	9. Definition of factoral
10. 1 < 2	10. Properties of the real numbers
11. 1 < 2 $\vee$ 1 = 2	11. Law of Addition (10)
12. 1 $\le$ 2	12. Definition of less than or equal to.
13. $\sum_{j=1}^{1} j(j!) \le (1+1)!$	13. Substitution
14. Assume $\exists$ m $\in$ $\mathbb{N}$ э $\sum_{j=1}^{m} j(j!) \le (m+1)!$	14. Inductive Hypothesis
15. Consider $\sum_{j=1}^{m+1} j(j!)$	15. Hypothesis
16. = 1(1!) + 2(2!) + ...+ m(m!) + (m+1)[(m+1)!]	16. Definition of Sigma
17. = {1(1!) + 2(2!) + ...+ m(m!)} + (m+1)[(m+1)!]	17. Associative of +

18. $= \sum_{j=1}^{m} j(j!) + (m+1)[(m+1)!]$	18. Definition of Sigma
19. $\leq (m+1)! + (m+1)[(m+1)!]$	19. Substitution
20. $= (m+1)!\,[1 + (m+1)]$	20. Distributive of $\times$ over +
21. $= (m+1)!\,[(m+1) + 1]$	21. Commutative of +
22. $= ((m+1)+1)!$	22. Definition of factoral
23. Thus, $\sum_{j=1}^{m+1} j(j!) = ((m+1)+1)!$	23. Transitivity of equality.

24. So, $\sum_{j=1}^{m} j(j!) \leq (m+1)! \Rightarrow \sum_{j=1}^{m+1} j(j!) \leq ((m+1)+1)!$

Q. E. D.

§ 3.6 EXERCISES.

Prove or disprove the following:

1. Claim: $\sum_{j=1}^{n} j^2 = \frac{n(n+1)(2n+1)}{6} \quad \forall \; n \in \mathbb{N}$

2. Claim: $2^n < 2^{(n+1)} \quad \forall \; n \in \mathbb{N}$

3. Claim: $2^n > n \quad \forall \; n \in \mathbb{N}$

4. Claim: $2^n < n! \quad \forall \; n \in \mathbb{N}$

5. Claim: $2^n \leq \prod_{j=1}^{n} j \quad \forall \; n \in \mathbb{N} - \mathbb{N}_3$

6. Claim: Let $a \in \mathbb{R} \; \wedge \; a > 1$. It is the case that $(1 + a)^n \geq 1 + an \quad \forall \; n \in \mathbb{N}$

7. Claim: $\sum_{j=1}^{n} (j \cdot (j+1))^{-1} = n \cdot (n+1)^{-1} \quad \forall \; n \in \mathbb{N}$

8. Claim: $\sum_{j=1}^{n} 2^j = 2^{(n+1)} - 2 \quad \forall \; n \in \mathbb{N}$

9. Claim: $2 + 4 + 6 + \ldots + 2n = n(n + 1) \quad \forall \; n \in \mathbb{N}$

10. Claim: $1 + 3 + 5 + \ldots + (2n - 1) = n^2 \quad \forall \; n \in \mathbb{N}$

11. Claim: $1 - 3 + 5 - 7 + 9 - 11 + \ldots + (-1)^n (2n - 1) = (-1)^n n^2 \quad \forall \; n \in \mathbb{N}$

12. Claim: $\sum_{j=1}^{n} 3^n = \frac{3^{(n+1)} - 3}{2} \quad \forall \; n \in \mathbb{N}$

13. Claim: Let $a \in \mathbb{R} \; \wedge \; b \in \mathbb{R}$. It is the case that $(a \cdot b)^n = a^n \cdot b^n \quad \forall \; n \in \mathbb{N}$

14. Claim: $\sum_{j=1}^{n} \binom{n+1}{2} = \binom{n+2}{3} \quad \forall \; n \in \mathbb{N}$

15. Claim: $\sum_{j=1}^{n} j \cdot (j+1) = \frac{n(n+1)(n+2)}{3} \quad \forall \; n \in \mathbb{N}$