II : Basic concepts

This unit is the beginning of the strictly mathematical development of set theory in the course. We begin with a brief discussion of how mathematics is written and continue with a summary of the main points in logic that arise in mathematics. The latter is mainly meant as background and review, and also as a reference for a few symbols that are frequently used as abbreviations. In the remaining sections we introduce the most essential notions of set theory and some of their simplest logical interrelationships.

Mathematical language

> Mathematicians are like Frenchmen; whatever you say to them they translate into their own language and forthwith it is something entirely different.
>
> J. W. von Goethe (1749 – 1832)

A page of mathematical writing is different from a page of everyday writing in many respects, and for an inexperienced or uninitiated reader it is often more difficult to understand. Before considering strictly mathematical topics in these notes, it might be helpful to summarize some special features of mathematical language and the reasons for such differences.

The language of mathematics is a special case of ***technical language*** or ***language for special purposes.*** As such, it has many things in common with other specialized language uses in the other sciences and also in legal writing.

In all these contexts, it is important to state things precisely and to justify assertions based upon earlier writing. It is also important to avoid things which are unrelated to the substance of the discussion, including emotional appeals and nearly all personal remarks; when the latter appear, they are usually restricted to a small part of the text.

The need for precise, impersonal language affects mathematical writing in several ways. We shall list some notable features below.

1. Sentences tend to be long and carefully written, sometimes at the expense of clarity. This is often necessary to avoid misunderstandings. For example, in mathematics when one divides a number **x** by a number **y,** it is necessary to stipulate that **y** be nonzero.
2. In scientific writing there is more of a tendency to stress nouns and modifiers rather than verbs, and there is a much greater use of the passive voice. For example, instead of saying, "You can do **X**," one generally sees the more

impersonal, "It is possible to do **X.**" This reinforces the unimportance or anonymity of the individual who does **X.** However, a reader who is not used to such an impersonal style might view it as uninviting.

3. Precise meanings must be attached to specific words. These do not necessarily correspond to a word's everyday meaning(s), and of course there are also many words that are rarely if ever seen elsewhere. Words like "product" and "set" and "differentiate" are examples of words whose mathematical meanings differ from standard usage. Other words such as "abelian" or "eigenvector" or "integrand" are essentially unique to mathematics and only appear when mathematics is presented or applied to another subject.
4. There is an extensive use of references to the writings of others. Such citations are logically indispensable and make everything more concise, but they can also make it difficult or impossible to read through something without frequent interruptions.
5. Particularly in the sciences, there is a heavy reliance on symbols such as numerals, operators (for example, the plus and equals signs), formulas or equations, and diagrams as well as other graphics. These allow the writer to express many things quickly but precisely. However, they may be difficult to decipher, particularly for a beginner.

The pros and cons of mathematical (and other scientific) language are reflected by a surprising fact: Even though such material is more difficult to read than an ordinary book, it is much easier to translate scientific writings to or from a foreign language than it is to translate a best seller or a regular column in a newspaper. In particular, adequate computerized translations of scientific articles are considerably easier to produce than acceptable computerized translations of literature.

Both clarity and preciseness are important in mathematical (and other scientific) writing. A lack of precision can lead to costly mistakes in scientific experiments and engineering projects (similar considerations apply to legal writing, where ambiguities involving simple words can lead to extensive and expensive litigation). On the other hand, a lack of clarity can undermine the fundamental goals of communicating information. Every subject has tried to adopt guidelines for balancing these contrasting aims, but probably there will always be challenges to doing so effectively in all cases.

II.0 : Topics from logic

(Lipschutz, §§ 10.1 – 10.12)

Mathematics is based upon logical principles, and therefore some understanding of logic is required to read and write mathematics correctly. In this course we shall take the most basic concepts of logic for granted. Our main purpose here is to describe the key logical points and symbolic logical notation that will be used more or less explicitly in this course. Chapter 10 of Lipschutz contains numerous examples illustrating the main points of logic that we shall use in this course, and it it provides additional background and reference material. Sections 1.1 – 1.5 of Rosen also treat these topics in an introductory but systematic manner.

In most mathematical writings, the logical arguments are carried out using ordinary language and standard algebraic symbolism. When logical terminology as developed in this section is used, it is often used intermittently for purposes of abbreviation when ordinary wording becomes too lengthy or awkward; there are similarities between this and the practice of explaining some programming issues in a pseudo – code that is halfway between ordinary and computer language. Although such logical abbreviations are only used sometimes in mathematics, it is important to be familiar with them and recognize them when they do appear.

Concepts from propositional calculus

The basic objects in propositional calculus are simple declarative sentences, and by convention each sentence is either true or false. There are several simple grammatical and logical operations that can be used to connect sentences.

1. If **P** and **Q** are sentences, then the sentence **P *and* Q** is sometimes called the **conjunction** of **P** and **Q**, and it is symbolically denoted by either $\mathbf{P} \wedge \mathbf{Q}$ or the less formal **P & Q**. Of course, if **P** and **Q** are both true, then $\mathbf{P} \wedge \mathbf{Q}$ is true, while if one or both of **P** and **Q** are false, then $\mathbf{P} \wedge \mathbf{Q}$ is false.
2. If **P** and **Q** are sentences, then the sentence **P *or* Q** is sometimes called the **disjunction** of **P** and **Q**, and it is denoted symbolically by $\mathbf{P} \vee \mathbf{Q}$**.** In mathematics we use an ***inclusive* OR *connective***; *i.e.*, $\mathbf{P} \vee \mathbf{Q}$ is true when **P** is true or **Q** is true, or both are true, and $\mathbf{P} \vee \mathbf{Q}$ is false only when both **P** and **Q** are false.
3. If **P** is a sentence, then the sentence ***not*** **P** is sometimes called the **negation** of **P,** and it is denoted symbolically by $\neg\, \mathbf{P}$ or **– P** or ***~P*** (still other symbolisms are also used). As one would expect, the sentence $\neg\, \mathbf{P}$ is false when **P** is true, and the sentence $\neg\, \mathbf{P}$ is true when **P** is false.
4. If ***P*** and ***Q*** are sentences, the **conditional** sentence ***if*** **P*, then*** **Q** is denoted symbolically by $\mathbf{P} \rightarrow \mathbf{Q}$ or $\mathbf{P} \Rightarrow \mathbf{Q}$**.** In the conditional sentence **P** is called the **antecedent** and **Q** is called the **consequent.** Such a conditional sentence is true unless **P** is true and **Q** is false, and it is false in this case. (The truth of the conditional statement if **P** is false may seem puzzling, but one way to think about it is that since **P** is false the conditional is basically an empty statement).

Of course, one can use the preceding connectives to define new ones in other ways, and one example is the ***exclusive* OR *connective*:** If **P** and **Q** are sentences, then the sentence **P *xor* Q** should have the property that **P *xor* Q** is false when **P** and **Q** are both true, and **P *xor* Q** is false otherwise. Symbolically one can write this connective in terms of the others by the formula $(\mathbf{P} \vee \mathbf{Q}) \wedge \neg\, (\mathbf{P} \wedge \mathbf{Q})$**.**

Another important operation is the standard ***if and only if connective.*** If ***P*** and ***Q*** are sentences, the **biconditional** sentence **P *if and only if* Q,** which is sometimes also written **P *iff* Q,** is simply **$(P \Rightarrow Q)$ & $(Q \Rightarrow P)$,** and it is written symbolically as **$P \Leftrightarrow Q$.** As expected, this statement is true if both **P** and **Q** are true or both are false, and it is false if exactly one of **P** and **Q** is true and exactly one is false**.** The phrase **P *is logically equivalent to* Q** is also used frequently in mathematical writings to denote the biconditional **$P \Leftrightarrow Q$.**

Tautologies

By definition, a ***tautology*** is a sentence that is true no matter what the truth values are for the constituent parts. One simple example of this is **$P \Rightarrow P \vee Q$.** Here are several others:

1. $(P \Rightarrow Q) \Leftrightarrow (\neg Q \Rightarrow \neg P)$ <u>*Law of the contrapositive*</u>
2. $[P \wedge (P \Rightarrow Q)] \Rightarrow Q$ <u>*Law of modus ponens*</u>
3. $[(P \Rightarrow Q) \wedge (Q \Rightarrow R)] \Rightarrow (P \Rightarrow R)$ <u>*Law of Syllogism*</u>
4. $\neg (P \wedge Q) \Leftrightarrow (\neg P \vee \neg Q)$
5. $\neg (P \vee Q) \Leftrightarrow (\neg P \wedge \neg Q)$ <u>*DeMorgan's Laws*</u>
6. $\neg (P \Rightarrow Q) \Leftrightarrow (P \wedge \neg Q)$
7. $(P \Rightarrow Q) \Leftrightarrow (\neg P \vee Q)$
8. $(P \wedge Q) \Rightarrow P$
9. $\neg (\neg P) \Leftrightarrow P$
10. $(P \wedge Q) \Rightarrow (P \vee Q)$
11. $(P \Rightarrow \neg Q) \Rightarrow (Q \Rightarrow \neg P)$
12. $[\neg P \Rightarrow (R \wedge \neg R)] \Rightarrow P$ <u>*Law of proof by contradiction*</u>
13. $[(P \wedge \neg Q) \wedge (R \wedge \neg R)] \Rightarrow (P \Rightarrow Q)$ <u>*Law of proof by contradiction*</u>
14. $P \wedge \neg P$ <u>*Law of the Excluded Middle*</u>
15. $P \Rightarrow P$
16. $P \Leftrightarrow P$
17. $[P \Rightarrow (Q \wedge R)] \Rightarrow [(P \wedge \sim Q) \Rightarrow R]$
18. $[(P \Rightarrow S_1) \wedge (S_1 \Rightarrow S_2) \wedge . . . \wedge (S_{n-1} \Rightarrow S_n) \wedge (S_n \Rightarrow R)] \Rightarrow (P \Rightarrow R)$
    <u>*Extended Law of Syllogism*</u>
19. $[(P \Rightarrow R) \wedge (Q \Rightarrow R)] \Rightarrow [(P \vee Q) \Rightarrow R]$ <u>*Proof by Cases*</u>
20. $(P \wedge Q) \Leftrightarrow (Q \wedge P)$
21. $(P \vee Q) \Leftrightarrow (Q \vee P)$ <u>*Commutative Laws*</u>
22. $[P \Rightarrow (R \Rightarrow Q)] \Leftrightarrow [(P \wedge R) \Rightarrow Q]$
23. $[P \wedge (Q \wedge R)] \Leftrightarrow [(P \wedge Q) \wedge R]$
24. $[P \vee (Q \vee R)] \Leftrightarrow [(P \vee Q) \vee R]$ <u>*Associative Laws*</u>
25. $[P \wedge (Q \vee R)] \Leftrightarrow [(P \wedge Q) \vee (P \wedge R)]$
26. $[P \vee (Q \vee R)] \Leftrightarrow [(P \vee Q) \wedge (P \vee R)]$ <u>*Distributive Laws*</u>
27. $[(P \Leftrightarrow Q_1) \wedge \ . . . \wedge (Q_{n-1} \Leftrightarrow Q_n) \wedge (Q_n \Leftrightarrow Q)] \Rightarrow (P \Leftrightarrow Q)$

Propositional calculus is covered in Sections 10.1 – 10.10 of Lipschutz and Sections 1.1 and 1.2 of Rosen. The material in these sections on the order of logical operations, translating English sentences and logic puzzles goes beyond the topics covered here.

Predicate calculus and quantifiers

Propositional calculus views sentences as units, and predicate calculus views ordinary declarative sentences as consisting of two main grammatical parts – the ***subject*** and the ***predicate.*** The subjects of such sentences are generally denoted by small letters like an **x**, and the predicates are denoted by functions like **P(...),** the idea being that given a predicate shell one can insert an arbitrary subject to obtain a grammatically admissible sentence **P(x)** which is either true or false. A typical example of such a sentence **P(x)** might be **x + 2 = 5.** For this example we know that **P(3)** is true but **P(4)** is false. Of course, ordinary sentences may have compound subjects, and it is essential to allow logical predicates to have this property also. As one might expect, we denote the sentence obtained from insertion of $\mathbf{x_1, \dots, x_n}$ into the predicate **P** by $\mathbf{P(x_1, \dots, x_n)}$**.**

We now turn to a discussion of **quantifiers**. Sentences involving the phrases like ***For every ...*** and ***There exists ...*** play a very important role in mathematically reasoning.

The logical symbol ∀, which is called the **universal quantifier**, is a symbolic shorthand for phrases such as For each, For every, For all. A predicate sentence such as ***For every* x, P(x)** is then written symbolically as either **∀x P(x)** or equivalently **∀x, P(x).** Here is a typical example of a true sentence in this form:

∀x, if **x** is a real number then $\mathbf{x^2}$ is nonnegative**.**

The logical symbol ∃, which is called the **existential quantifier**, is a symbolic shorthand for phrases such as There exists, There is at least one, For at least one, and For some. A sentence such as ***There exists an* x *such that* P(x)** is then written symbolically as either **∃ x P(x)** or equivalently **∃ x, P(x).** Here is a typical example of a true sentence in this form:

∃ x, if **x** is a real number then 1 **–** $\mathbf{x^2}$ is nonnegative**.**

Note that if **P** is the predicate in the sentence above, then **∃ x, P(x)** is true (take **x = ½**) but **∀x, P(x)** is false (take **x = 2**). On the other hand, for every predicate **Q** we know that **∀x Q(x)** automatically implies **∃ x Q(x).**

Since we are discussing tautologies involving quantifiers, we should mention two other basic statements of this type.

Tautology Criterion 1: *Every sentence of the type* **[¬∃x, P(x)] ⇔ [∀x, ¬P(x)]** *is true.*

Tautology Criterion 2: *Every sentence of the type* **[¬∀x, P(x)] ⇔ [∃x, ¬P(x)]** *is true.*

In mathematical writings one often sees a variant of the existential quantifier called the ***unique existential quantifier***, which is denoted by **∃|** or **∃!** or **∃1** and signifies the ***unique existence*** of some object. For example, the sentence **∃| x, P(x)** is true if **P(x)** is given as follows:

$\mathbf{x}$ is an integer and $\mathbf{x + 1 = 2.}$

Formally one can express ∃| directly in terms of the other quantifiers because a statement of the form ∃| **x, P(x)** can be written in the following equivalent terms:

$$[\, \exists\, x, P(x)\,] \;\;\&\;\; [\, \forall x\, \forall y, \{\, P(x)\ \&\ P(y)\,\} \Rightarrow \{\, x = y\,\}\,]$$

Another point about quantifiers that merits discussion is the order in which they are listed. If an expression contains multiple quantifiers, the order in which they appear may be very important. For example, suppose that **P(x, y)** is the following statement:

$\mathbf{x}$ is a real number, and if $\mathbf{y}$ is a real number then $\mathbf{x > y.}$

Then ∀**y** ∃**x, P(x, y)** means that for every real number **x** there is a larger real number **y,** and hence the quantified statement is ***true*** **,** but ∃**x** ∀**y, P(x, y)** is ***false*** (there is no number **x** which is greater than every number, including itself). In contrast, if **P** is a predicate such that ∃**x** ∀**y, P(x, y)** is true, then ∀**y** ∃**x, P(x, y)** will *always* be true.

Predicate calculus is covered in Section 10.11 of Lipschutz and Sections 1.3 and 1.4 of Rosen. The material in these sections on bound variables, nested quantifiers, the order of quantifiers, translating English sentences and Lewis Carroll's logical puzzles goes beyond the topics covered here and in this course.

Formal structure of languages

The predicate calculus is an important first step in studying the formal structure or syntax of the language needed to carry out logical processes. The study of such structure is particularly important in some aspects of computer science. A detailed discussion of this topic is beyond the scope of these notes, but a good introductory discussion appears in Section 11.1 of Rosen. It is extremely interesting to note that much of the work on formal grammars by noted workers in computer science such as J. Backus (1924 –) was anticipated centuries ago in the profound analysis of Sanskrit grammar due to Panini (520 – 460 B. C. E.) in his ***Astadhyayi*** (or ***Astaka***)**.** It is particularly noteworthy that Panini's notation is equivalent in its power to that of Backus, and it has many similar properties.

Mathematical proofs

Standard methods and strategies for mathematical proofs are discussed in Sections 1.5, 3.1 and 3.3 of Rosen. We shall summarize the main points from these sections, mention a few other points points not specifically covered in these citations, and give some examples from high school mathematics and calculus (we are simply trying to illustrate the techniques, so our setting for now is informal, and in particular for the time being we shall not worry about things like how one proves the Intermediate Value Theorem that plays such an important role in calculus. This is technically an example of a concept called ***local deduction*****,** in which one only shows how to get from point **A** to point **B,** postponing questions about reaching point **A** to another time or place**.**

Some proofs use ***direct arguments***, while others use ***indirect arguments.*** The direct arguments are often the simplest, and many simple problem solving methods from elementary mathematics (algebra, in particular) are really just simple examples of direct proofs.

Example. If $\mathbf{2x + 1 = 5,}$ show that $\mathbf{x = 4.}$ **SOLUTION:** If $\mathbf{2x + 1 = 5,}$ then by subtracting 1 from each side we obtain $\mathbf{2x = 4.}$ Next, if we divide both sides of $\mathbf{2x = 4}$ by **2,** we obtain $\mathbf{x = 2.}$

In contrast, and indirect argument usually involves considering the negation of either the hypothesis or the conclusion. This generally involves ***proof by contradiction***, in which one assumes the conclusion is false and then proves part of the hypothesis is false, and it is related to the law of the ***contrapositive*****:** A statement $\mathbf{P \Rightarrow Q}$ is true if and only if the *contrapositive* statement $\mathbf{not\ Q \Rightarrow not\ P}$ is true.

A general **"**rule of thumb**"** is to consider using an indirect argument if either no way of using a direct argument is apparent or if a direct approach seems to be getting very long and complicated. There is no guarantee that an indirect argument will be any better, but if you get stuck trying a direct approach there often is not much to lose by seeing what happens if you try an indirect approach; in some cases, attempts to give an indirect argument may even lead to a valid or better direct proof.

Example. Show that if **L** and **M** are two lines then they have at most one point in common. **SOLUTION:** Suppose the conclusion is false, so that **x** and **y** are two distinct points on both **L** and **M.** Then both **L** and **M** are lines containing these two points. Since there is only one line **N** containing the two distinct points **x** and **y,** we know that **L** must be equal to **N** and similarly **M** must be equal to **N,** which means that **L** and **M** must be equal. This contradicts our original assumption; the problem arose because we added an assumption that **x** and **y** belonged to both lines. Therefore **L** and **M** cannot have two (or more) points in common.

An important step in such indirect arguments is to make sure that the negation of the conclusion is accurately stated. Mistakes in stating the negation usually lead to mistakes in arguments intended to prove the original result.

Forward and backwards reasoning. Very often it is helpful to work backwards as well as forwards. For example, if you want to show that **P** implies **Q,** in some cases it might be easier to find some statement **R** that implies **Q,** and then to see if it is possible to prove that **P** implies **R.** Of course, there may be several intermediate steps of this type.

Example. Show that the polynomial $\mathbf{f(x) = x^5 - x - 1}$ has a real root. **SOLUTION:** We know that polynomials are continuous and that continuous functions have the Intermediate Value Property. Therefore if we can show that the polynomial is positive for some value of **x** and negative for another, then we can also show that this polynomial has a real root. One way of doing this is simply to calculate the value of the polynomial for several different values of the independent variable. If we do so, then we see that $\mathbf{f(0) = -1}$ and $\mathbf{f(2) = 29.}$ Therefore we know that $\mathbf{f(x)}$ has a root, and in fact by the Intermediate Value Theorem from first year calculus we know there is a root which lies somewhere between **0** and **2.**

Proofs by cases. Frequently it is convenient to break things up into all the different cases and to check them individually, and in some cases this is simply unavoidable.

Example. Let $\mathbf{sgn(x)}$ be the function whose value is $\mathbf{1}$ if $\mathbf{x}$ is positive, $\mathbf{-1}$ if $\mathbf{x}$ is negative, and $\mathbf{0}$ if $\mathbf{x = 0}$. Prove that $\mathbf{sgn(xy) = sgn(x)\,sgn(y)}$. Then there are three possibilities for $\mathbf{x}$ (positive, negative, zero) and likewise for $\mathbf{y}$, leading to the following list of nine possibilities for $\mathbf{x}$ and $\mathbf{y}$:

$$[+,+],\ [+,0],\ [+,-],\ [0,+],\ [0,0],\ [0,-],\ [-,+],\ [-,0],\ [-,-]$$

One can then handle each case (or various classes of cases) separately; for example, the five cases where at least one number is zero follow because in all these cases we have $\mathbf{xy = sgn(x)\,sgn(y) = 0}$.

In all proofs by cases**, *it is important to be absolutely certain that all possibilities have been listed.*** The omission of some cases is an automatic mistake in any proof.

Interchanging roles of variables. This is a basic example of proofs by cases in which it is possible to **"**leverage**"** one case and obtain the other with little or no additional work.

Example. Show that if $\mathbf{x}$ and $\mathbf{y}$ have opposite signs, then we have $\mathbf{|x - y| = |x| + |y|}$.
SOLUTION: Suppose first that $\mathbf{x}$ is positive and $\mathbf{y}$ is negative. Then the left hand side is just $\mathbf{x + |y| = |x| + |y|}$. Now suppose $\mathbf{y}$ is positive and $\mathbf{x}$ is negative. Then if we apply the preceding argument to $\mathbf{y}$ and $\mathbf{x}$ rather than to $\mathbf{x}$ and $\mathbf{y}$ we then obtain the equation $\mathbf{|y - x| = |y| + |x|}$. Since the left hand side is equal to $\mathbf{|x - y|}$ and the right hand side is equal to $\mathbf{|x| + |y|}$, we get the same conclusion as before. In a situation of this type we often say that the second case follows from the first by reversing the roles of $\mathbf{x}$ and $\mathbf{y}$.

Vacuous proofs. In some instances a statement is true because there are no examples where the hypothesis is valid.

Example. Show that if $\mathbf{x}$ is a number such that $\mathbf{x + 1 = x}$, then $\mathbf{x^2 + 1 = x^2}$.
SOLUTION: There is no number satisfying the hypothesis, so whatever conclusion one states, there will be no number which satisfies the first but does not satisfy the second. Formally, the statement $\mathbf{P \Rightarrow Q}$ merely signifies that there are no situations in which $\mathbf{P}$ is true but $\mathbf{Q}$ is false; if there are no situations where $\mathbf{P}$ is true, then there also cannot be any where $\mathbf{P}$ is true but $\mathbf{Q}$ is false.

How can this be useful in mathematics? Sometimes the use of vacuously true statements allows one to state conclusions in a simpler or more uniform manner. For example, in elementary geometry one can show that the sum of the measures of the vertex angles for a regular $\mathbf{n}$ – gon is equal to $\mathbf{180(n - 2)/n}$ degrees**.** In some sense this is only valid if $\mathbf{n}$ is at least $\mathbf{3}$ because every regular polygon has at least three sides, but for some purposes it is convenient simply to state the formula for all positive integers $\mathbf{n}$. The formula gives a negative angle measurement for $\mathbf{n = 1}$, but this does not matter because this case of the formula does not apply if $\mathbf{n = 1}$ since there are no $\mathbf{1}$ – gons. The point is that the statement of the formula is logically correct even if we omit the condition that $\mathbf{n}$ is at least $\mathbf{3}$. This is a simple situation, but the concept of **"**vacuously true**"** also turns out to be useful in other situations where the hypothesis or conclusion is more complicated.

Adapting existing proofs. In all activities, it can be useful to use an idea that has worked to solve one problem in an attempt to solve another that may be somehow related. The same principle works for mathematical proofs. You can try this approach in order to prove that if $3x + 1 = 10$, then $x = 3$ (modify the first proof above).

Disproving conjectures. Frequently one is faced with an unproven statement and the goal is to determine whether it is true or false. If you suspect the statement is false, often the fastest way to confirm this is to construct a ***counterexample*** which satisfies the hypotheses but not the entire conclusion.

Illustration. If we are given real numbers a and b such that $a^3 - a = b^3 - b$, can we conclude that $a = b$? **SOLUTION:** We should remark first that this is true if the absolute values of a and b are greater than 2, and someone who knows this might wonder if it is evidence that the result is always true. However, it is not; to show this we need to find explicit distinct values of a and b for which the equation holds. This can be done systematically, but the fastest way is to look at some examples and notices that the numbers 0 and 1 provide a counterexample.

On the other hand, ***it is important to recognize that one cannot prove a general statement by simply checking one, or even infinitely many examples that do not exhaust all the possibilities,*** and the preceding statement demonstrates this very convincingly (it is true whenever a and b are greater than 2).

Contrapositives, biconditionals and logical equivalences. In order to complete a proof of the biconditional (or logical equivalence) statement $P \Leftrightarrow Q$, it suffices to prove the two separate statements $P \Rightarrow Q$ and (its "inverse" statement) **not** $P \Rightarrow$ **not** Q. [The reason for this rule is that the inverse statement **not** $P \Rightarrow$ **not** Q is the contrapositive of the converse statement $Q \Rightarrow P$.]

Similarly, in order to complete a proof of $P \Leftrightarrow Q$, it suffices to prove the contrapositive statement **not** $Q \Rightarrow$ **not** P and the inverse statement **not** $P \Rightarrow$ **not** Q.

Proofs of existence and uniqueness. It is *absolutely essential* to remember that *all such proofs have two parts*, one of which is an *existence proof* and the other of which is a *uniqueness proof*.

A symbolic approach to proofs. If it is difficult to decide how to start a proof, one suggestion is to put things into symbolic terms along the lines of the present section. This may provide enough insight into the question that a successful proof strategy can be found.

The use of definitions as a proof strategy. Another suggestion for finding a proof strategy is to recall all relevant definitions; it is very easy to overlook these or recall them inaccurately.

The do – something approach to finding proofs. This is simply trial and error, but it definitely should not be underestimated (recall Thomas Edison's comment about genius being 99 per cent perspiration and one per cent inspiration**!**). Even if no particular way of getting from the start to the finish is apparent, there is often little to lose by simply getting involved, doing something, trying different approaches, drawing pictures and proving everything that one can from the information given. Most of the proofs in print give no idea of the dead ends, incomplete arguments and otherwise unsuccessful efforts at proving something that took place before a valid proof was found. Trial and error is just as much a part of proofs in mathematics as it is of any other intellectual activity**.**

Mathematical induction (Finite induction). This is often a very powerful technique, but it is really more of a method to provide a formal verification of something that is suspected to be true rather than a tool for making intuitive discoveries, but it is absolutely essential. The use of mathematical induction dates back at least to some work of F. Maurolico (1494 – 1575). There are many situations in discrete mathematics where this method is absolutely essential; we shall postpone discussing this until Unit **V.**

Avoiding and finding mistakes in proofs. Unfortunately, there is no simple way of doing these outside of checking things repeatedly and carefully, but we have already mentioned a few common causes of difficulties and how to prevent them and there are several more common errors that can be mentioned: The list below is by no means exhaustive.

1. ***Begging the question.*** Frequently one finds arguments in which a proof uses and relies upon some other auxiliary which has not been proven. In such instances all one has shown is that *if this auxiliary statement is true, then the original statement is true.* However, we may have no way of knowing whether the auxiliary statement is true or false.

2. ***Computational errors.*** Sometimes mistakes in arithmetic or algebra are embedded in arguments and destroy their validity.

3. ***Incorrect citations of other results.*** Of course, this can be deadly to a proof. Division by zero is a standard elementary example in which one neglects to recognize that **ax = ay** *implies* **x = y** *only if* **a** *is nonzero.*

4. ***Proving only half of biconditional or existence – uniqueness proofs.*** Half a proof may be better than none at all, but it is still just half a proof.

5. ***Proving the converse instead.*** Often one finds arguments which show that if the conclusion is true, then the hypothesis is true. This is the reverse of what is supposed to be established.

6. ***Using unproven converses.*** This is a special case of the third item, but it is also one which plays a role in elementary algebra.

The last of these is related to material on ***extraneous roots*** that one finds in elementary algebra courses. Here is a quick review of the underlying ideas. Suppose that we want to solve an equation like

$$x - 3 = \sqrt{30 - 2x}$$

The standard way to attack this problem is to eliminate the radical by squaring both sides and solving for **x:**

$$(x-3)^2 = \left(\sqrt{30-2x}\right)^2$$
$$x^2 - 6x + 9 = 30 - 2x$$
$$x^2 - 4x - 21 = 0$$
$$(x-7)(x+3) = 0$$
$$x = 7; x = -3$$

(Source: http://regentsprep.org/Regents/mathb/7D3/radlesson.htm)

This tells us that the only possible solutions are given by the two values above, but ***it does not guarantee that either is a solution.*** The reason for this is that the first step, in which we square both sides, shows that the first equation implies the second, but it does not imply that the second implies the first; for example even though the squares of **2** and **– 2** are equal, it clearly does **not** follow that these two numbers are the same. In order to complete the solution of the problem, we need to go back and determine which, if any, of these two possible solutions will work. It turns out that **x = 7** is a solution, but on the other hand **x = – 3** is not (and hence is an extraneous root).

The online site http://www.jimloy.com/algebra/square.htm discusses further examples of this type.

Pólya's suggestions for solving problems. The classic book, ***How to solve it,*** by G. Pólya (1887 – 1985), discusses useful strategies for working problems in mathematics. A summary of his suggestions and a more detailed reference for the book appear in the online document

http://math.ucr.edu/~res/polya.pdf

which is stored in the course directory.

Ends of proofs. In classical writings mathematicians used the initials **Q. E. D**. (for the Latin phrase, *that which was to be demonstrated*) or **Q. E. F.** (for the Latin phrase, *that which was to be constructed*) to indicate the end of a proof or construction. Some writers still use this notation, but more often the end of a proof or line of reasoning is now indicated by a large black square, which is sometimes known as a "tombstone" or "Halmos (big) dot." We shall also use the symbol "■" to mark the end of an argument.

Reference for further reading. There is an article on writing proofs ("A guide to proof – writing," by R. Morash) on pages 437 – 447 of the following supplement to Rosen's text:

> K. Rosen, ***Student Solutions Guide to Discrete Mathematics and Its Applications*** (5^{th} Ed.). McGraw – Hill, Boston, 2003. ISBN: 0–07–247477–7.

Of course, there are also many other excellent books available; we have chosen one that is closely related to a text that was consulted repeatedly in the preparation of these notes.

II.1 : Notation and first steps

(Halmos, § 1; Lipschutz, §§ 1.2 – 1.5, 1.10)

We shall start by presenting the naïve approach, and then we shall explain how things can be set up more formally. A reader who wishes to skip the latter may do so by going directly from the end of the discussion of the former to the final portion of this section titled *A few simple consequences*.

The naïve approach

Most if not all of this is probably familiar, but it is necessary to state things explicitly for the sake of completeness.

In the mathematical sciences, a **"*set*"** is supposed to be a ***collection of objects*;** as noted on page 4 of Halmos, **"**A pack of wolves, a bunch of grapes or a flock of pigeons are all examples.**"** To illustrate the generality of the concept, we note that the objects in a set may themselves be sets. For mathematical purposes the only relevant information about a set concerns the objects belonging to it, and accordingly ***a set is completely determined by the objects that belong to (or are members of) it.*** If an object $\mathbf{x}$ belongs to a set $\mathbf{X}$, we shall denote this fact by the usual notation $\mathbf{x} \in \mathbf{X}$.

There are two standard ways of describing a set. In some cases we can describe the set by ***listing*** all the objects in it. For example, the set consisting of the positive integers from **1** to **5** may be denoted by $\{ 1, 2, 3, 4, 5 \}$**.** On the other hand, a set is often described in terms of the properties that are true for objects belonging to it and false for objects that do not belong to it. For example, if we wish to describe the set of whole numbers that are perfect squares, we use what is called ***set builder notation*:**

$$\{ \mathbf{x} \mid \mathbf{x} \text{ is an integer and } \mathbf{x} = \mathbf{y}^2 \text{ for some integer } \mathbf{y} \}$$

This is read verbally as **"**the set of all $\mathbf{x}$ such that $\mathbf{x}$ is an integer and $\mathbf{x}$ is equal to $\mathbf{y}^2$ for some integer $\mathbf{y}$**"** (where the vertical line **"|"** is read **"**such that**"**).

The possibility of a set which has no members is generally allowed, and it is called the **"*empty set*"** (or ***null set***). It is generally denoted by symbolism such as **Ø.**

A **"*subset*"** of a set $\mathbf{X}$ is simply a set which contains some but not necessarily all of the objects in $\mathbf{X}$, and it is a "**proper subset**" if it does not contain all of the objects in $\mathbf{X}$. Subsets are denoted using the symbol $\subset$, and the statement $\mathbf{Y} \subset \mathbf{X}$ is often expressed verbally as **"Y** *is a subset of* **X"** or **"Y** *is contained in* **X"** or **"X** *contains* **Y."** Sometimes we shall also express this relationship using the notation $\mathbf{Y} \supset \mathbf{X}$.

There is one further point which is usually omitted in elementary treatments of set theory but must be mentioned here. Although there is a great deal of flexibility in the sorts of properties that can be used to define a set, serious problems arise if one tries to stretch this too far. Such difficulties were first discovered at the end of the 19^{th} and beginning of the 20^{th} century and involve collections that are somehow **"**too big**"** to be handled effectively. For example, problems arise if one tries to talk about **"**the set of all possible sets.**"** Further information on this appears on pages 6 – 7 of Halmos and in the more formal approach to set theory in this section.

There are two ways of avoiding such problems with oversize collections. One is to recognize their existence but to have a two-tiered system of collections in which some are regarded as sets and others are not. The latter are generally too large, and one cannot do as much with them as one can with sets. For example, a collection which is not a set cannot be viewed as a member of some other collection. Fortunately, these exceptional objects do not cause any real problems most of the time; in nearly all situations, the foundational questions can be avoided by assuming that everything in sight lies inside some very large and fixed <u>*quasi – universal*</u> set.

Once again, a reader who wishes to skip the more formal discussion of the framework for set theory may do so by proceeding directly to the heading, *A few simple consequences*.

A more formal approach

> Nothing will come of nothing.
>
> (Shakespeare, ***King Lear,*** Act **I**, Sc. 1)

Every logical discussion must begin somewhere. An endless sequence of definitions or proofs based on earlier ones will not lead to any firm conclusions. In order to begin, the following three requirements must be fulfilled:

1. There must be a mutual understanding of the words and symbols to be used.
2. There must be acceptance of certain statements whose correctness is not further justified.
3. There must be agreement about the ***rules of reasoning*** which determine how and when one statement follows logically from another.

The words and symbols in the first item are generally known as ***undefined concepts*** in mathematics, and the statements described in the second item are generally known as ***assumptions, axioms*** or ***postulates*** (in modern usage all these are synonymous). We have already treated the rules of reasoning in Section **II.0.**

By modern standards, one logical difficulty with Euclid's ***Elements*** is that it tried to define everything. For example, a point was defined to be something that had no **"**part**"** or dimensions; to be logically precise, such a definition depends in turn upon giving a sound definition of "part" or dimension, and of course the same applies to any terms used in the definitions of the latter. The introduction of undefined concepts eliminates such infinite regressions. However, it is important to recognize that undefined concepts may not have any real value unless one has some understanding of what they are supposed to represent. In other words, if deductions are expected to yield useful information, then the undefined concepts in a discussion should be formal idealizations of things that are relatively familiar and recognizable.

Undefined concepts in set theory

Not surprisingly, the most important undefined concept in this subject is a set, which corresponds to a collection of objects. Since one important property of such a collection is whether some given object belongs to it, the notion of one entity belonging to another is almost as basic of an undefined concept as the notion of a set itself.

In order to avoid logical difficulties with oversized sets described above, we shall work with ***three*** primitive concepts which reflect the intuitive notions in the preceding paragraph.

1. ***CLASSES.*** These are collections of objects; it is assumed that each object itself is also a class.
2. ***SETS.*** Collections of objects that are small enough to work with reliably.
3. ***MEMBERSHIP.*** A grammatical statement with two subjects that represents one class belonging to another.

Items of the first type (actually, two types) are generally denoted by symbols such as letters. The statement that a class **A** belongs to a class **B** is usually written in the standard manner as $\mathbf{A} \in \mathbf{B}$. Likewise, we shall write $\mathbf{A} \notin \mathbf{B}$ to indicate that **A** does ***NOT*** belong to the class **B.** Following standard mathematical usage, we shall often use expressions of the following types as synonyms for $\mathbf{A} \in \mathbf{B}$:

- **A** <u>*belongs to*</u> **B.**
- **A** <u>*is a member of*</u> **B.**
- **A** <u>*is an element of*</u> **B.**

Furthermore, we shall often say that the <u>*members*</u> or <u>*elements*</u> of a class **B** are all the objects **A** such that $\mathbf{A} \in \mathbf{B}$. None of this is surprising, but the important point is that we are trying to build a theory of sets that is completely formal starting from scratch, and we need to start with this familiar sort of structure.

<u>*Comments on the introduction of classes as an undefined concept.*</u> Our approach, which differs from Halmos in that we also mention certain collections of objects that are too large to be treated as sets; was developed by J. von Neumann (1903 – 1957). As an example of the logical problems with an overly casual approach to set theory that are discussed in pages Halmos, we note that difficulties arise if one attempts to consider a ***universal set*** containing all sets. More will be said about this in the discussion of Russell's Paradox in Section **II.3.** The viewpoint of these notes resembles the approach taken in many versions of axiomatic set theory: It ***is*** meaningful for us to talk

about a universal collection or ***class*** of objects, but the latter is simply too large to be treated as a set. If a class is ***NOT*** a set, we shall say that it is a ***proper class.***

Our first basic assumption will be a smallness property that characterizes sets.

SMALLNESS PROPERTY FOR SETS. *A class* **A** *is a set if and only if* **A** $\in$ **B** *for some class* **B.**

Some good news. As we have already noted, in mathematics it is usually not necessary to worry very much about the formal distinction between sets and classes. The following paragraph summarizes the situation:

> For all practical purposes within this course, and nearly all other purposes in higher mathematics, one can simply view a set as a collection of objects that is not too large; a standard way of doing this is to ***assume that all objects in a given situation are subsets of some fixed larger set.***

The most significant exceptions to this principle arise in material dealing explicitly with the foundations of mathematics.

The definitions of *subclass* and *subset* are now straightforward.

Definition. Let **A** and **B** be classes of objects. We shall say that **A** is a *subclass* of **B** and write **A** $\subset$ **B** if for each object **x** such that **a** $\in$ **A,** then we also have **x** $\in$ **B.** If in addition **A** and **B** are sets, then we shall say that **A** is a ***subset*** of **B**.

If **A** $\subset$ **B** and the class **B** is small enough to be a set then one would expect the same holds for the class **A,** and in fact this is the case.

SUBSET PROPERTY. *If* **A** $\subset$ **B** *and* **B** *is a set, then* **A** *is also a set.*

Previous experience with set theory suggests that two sets should be the same if and only if they contain exactly the same objects. The next property reflects this basic fact.

EXTENSIONALITY PROPERTY. *If* **A** *and* **B** *are classes, then* **A = B** *if and only if we have* **A** $\subset$ **B** *and* **B** $\subset$ **A.**

Finally, we need to add another simple assumption, without which the whole theory would be entirely meaningless.

MINIMAL EXISTENCE PROPERTY. *There exists at least one set.*

A few simple consequences

Regardless of whether we adopt a naïve or more formal approach to set theory, there are already a few conclusions that can derived from what we have developed thus far.

Here are two simple but important logical consequences of the definition of a subset or subclass:

<u>Proposition 1.</u> *For each class* **A** *we have* $\mathbf{A} \subset \mathbf{A}$.

<u>Proof.</u> By definition of subclasses, this amounts to saying that for all $\mathbf{x}$ such that $\mathbf{x} \in \mathbf{A}$, we have $\mathbf{x} \in \mathbf{A}$. But this follows because every true statement implies itself.■

<u>Definition.</u> If **A** and **B** are classes of objects such that $\mathbf{A} \subset \mathbf{B}$, we shall say that **A** is a <u>proper subclass</u> of **B** if in addition $\mathbf{A} \neq \mathbf{B}$ (and a <u>proper subset</u> if **B** is a set)**.**

<u>Proposition 2.</u> *If we are given classes* **A, B, C** *such that* $\mathbf{A} \subset \mathbf{B}$ *and* $\mathbf{B} \subset \mathbf{C}$, *then we also have* $\mathbf{A} \subset \mathbf{C}$.

<u>Proof.</u> By definition of subclasses and the assumptions, we know that for each $\mathbf{x}$ such that $\mathbf{x} \in \mathbf{A}$, we also have $\mathbf{x} \in \mathbf{B}$. Likewise, for each $\mathbf{y}$ such that $\mathbf{y} \in \mathbf{B}$, we also have $\mathbf{y} \in \mathbf{C}$. Combining these, we conclude that for each $\mathbf{x}$ such that $\mathbf{x} \in \mathbf{A}$, we must also have $\mathbf{x} \in \mathbf{C}$.■

The Extensionality Property (two classes are the same if they have the same elements) has a simple but fundamental consequence.

<u>Proposition 3.</u> *If* **A** *is a* ***proper*** *subclass of* **B,** *then there exists some object* **x** *such that* $\mathbf{x} \in \mathbf{B}$ *but* $\mathbf{x} \notin \mathbf{A}$.

<u>Proof.</u> By hypothesis we know that $\mathbf{A} \subset \mathbf{B}$ but $\mathbf{A} \neq \mathbf{B}$. If $\mathbf{B} \subset \mathbf{A}$ were true, then by extensionality we would have $\mathbf{A} = \mathbf{B}$. Therefore $\mathbf{B} \subset \mathbf{A}$ must be false, and this means that there must be some $\mathbf{x}$ such that $\mathbf{x} \in \mathbf{B}$ but $\mathbf{x} \notin \mathbf{A}$.■

Variants of sets

For certain purposes it is useful to have elaborations of sets known as <u>*multisets*</u> (also called ***bags***) and <u>*fuzzy sets*</u>. For both of these, the extra data are numerical “values of membership” attached to each element. In the case of multisets, the value is a positive integer and it indicates that an element is somehow repeated; a simple example would be the roots of a quadratic equation, where one might have two single roots or one double root. For fuzzy sets, the value of membership is a real number in the unit interval, and intuitively it can be viewed as a probability that the element actually belongs to the set in question. Further discussions of both concepts are given on pages 96 – 97 of Rosen.

II.2 : Simple examples

(Halmos, §§ 1 – 3; Lipschutz, § 1.12)

Thus far the only specific example of a set we have mentioned is the empty set, and at this point we need some ways of constructing other examples. Once again, prior experience with set theory suggests that one can define a set by stipulating that the objects contained in it satisfy a given condition. Our next order of business is to make this more precise; the following version covers both the naïve and formal approaches..

<u>**SPECIFICATION PROPERTY.**</u> *Suppose that we are given a set* **A** *and an admissible predicate statement* **P(x).** *Then there is a subset* $\mathbf{B} \subset \mathbf{A}$ *such that* $\mathbf{x} \in \mathbf{B}$ *if and only if* $\mathbf{x} \in \mathbf{A}$ *and the statement* **P(x)** *is true***.**

To elaborate on comments in the previous section, some standard ways of writing such a set are

$$\{\, \mathbf{x} \mid \mathbf{x} \in \mathbf{A}\ \&\ \mathbf{P(x)} \,\} \quad \text{or} \quad \{\, \mathbf{x} \in \mathbf{A} \mid \mathbf{P(x)} \,\} \quad \text{or} \quad \{\, \mathbf{x} \in \mathbf{A} : \mathbf{P(x)} \,\}.$$

The admissibility requirement is included to guarantee that the statement **P(x)** is meaningful in our context; for most practical purposes it will create no problems. A brief discussion of suitably meaningful statements appears on pages 5 – 6 of Halmos.

It is possible to weaken the Specification Axiom somewhat to eliminate the dependence on some predetermined set **A,** but in practice this requirement is not an obstacle and the weaker statement is considerably more complicated to state. However, some additional condition is needed to avoid logical difficulties. We shall not give an explicit description of admissibility, but it is useful to discuss the problems which shoed the need for such a restriction/

Admissible statements and Russell's Paradox

The most convincing example to illustrate the need to avoid totally unrestricted constructions of the form

$$\{\, \mathbf{x} \mid \mathbf{P(x)} \,\}$$

was discovered by B. Russell (1872 – 1970; known outside of mathematics for his philosophical writings and political activism) near the beginning of the 20th century. He considered the simple example where **P(x)** is given by $\mathbf{x} \notin \mathbf{x}$. Suppose we can construct a set $\mathbf{A} = \{\, \mathbf{x} \mid \mathbf{x} \notin \mathbf{x} \,\}$. One can then ask whether or not $\mathbf{A} \in \mathbf{A}$. If the answer is ***<u>yes</u>*****,** then the definition of **A** would seem to imply that $\mathbf{A} \notin \mathbf{A}$, while if the answer is ***<u>no</u>*****,** then the definition of **A** would seem to imply that $\mathbf{A} \in \mathbf{A}$. Each options leads to a contradiction, and hence neither is acceptable. Numerous other problems of a similar nature were discovered around the same time. Eventually it became clear that the underlying difficulty resulted from attempts to use sentences which somehow refer to themselves (think about the nonmathematical statements, "This sentence is false," or "All generalizations are incorrect."). The specific condition in our Specification Axiom is a simple but effective way of doing so.

The idea of a set being an element of itself is somewhat contrary to our intuition, and in the usual forms of set theory in use today the possibility is excluded. We shall discuss this further in Section **III.5.**

A formal approach to the empty set

A reader who wishes to bypass material on the formal approach to set theory may skip this discussion and proceed directly to the next heading.

We have not yet explained how or why the empty set fits into our formal approach to set theory. The Specification Property gives us an easy way of doing so.

Proposition 1. *In the formal approach to set theory, there is a unique **empty set*** $\emptyset$ *with the property that* $\mathbf{x} \notin \emptyset$ *for every set* **x.**

Proof. Since we just assumed the existence of a set, let us try to use this right away. Let A be a set, and use the Specification Axiom to construct the set

$$\mathbf{N} = \{ \mathbf{x} \in \mathbf{A} \mid \mathbf{x} \neq \mathbf{x} \}.$$

For all $\mathbf{y} \in \mathbf{A}$ we have $\mathbf{y} = \mathbf{y}$ and therefore it follows that $\mathbf{y} \notin \mathbf{N}$ for all $\mathbf{y} \in \mathbf{A}$. By construction it follows that $\mathbf{z} \notin \mathbf{N}$ for all $\mathbf{z} \notin \mathbf{A}$, and therefore we conclude that $\mathbf{x} \notin \mathbf{N}$ for every set **x.** This proves the existence part of the proposition.

To prove uniqueness, let **M** and **N** be sets such that $\mathbf{x} \notin \mathbf{M}, \mathbf{N}$ for every set **x.** Since nothing belongs to either set we trivially have $\mathbf{M} \subset \mathbf{N}$ and $\mathbf{N} \subset \mathbf{M}$, and therefore by the Extensionality Property we must have $\mathbf{M} = \mathbf{N}$.■

Important special cases of the Specification Axiom

Informally the uses of the specification axiom to construct sets should be clear. For example, if we have a set R of real numbers with the expected properties then we can define the ***closed interval***

$$[0, 1] = \{ \mathbf{x} \in \mathbf{R} \mid 0 \leq \mathbf{x} \leq 1 \}$$

and similar subsets that arise repeatedly in calculus and other mathematics courses. Our interest here will be more directed towards simple general constructions. The remainder of this section is valied for both the naïve and formal approaches.

Proposition 2. *Suppose that* $\mathbf{A}$ *is a set. Then there is a set* $\mathbf{\{ A \}}$ *such that* $\mathbf{A \in \{ A \}}$ *if and only if* $\mathbf{x = A}$.

Proof. Since $\mathbf{A}$ is a set we know that $\mathbf{A \in B}$ for some $\mathbf{B}$. By the Specification Axiom there is a set given by the description $\mathbf{\{ x \in B \mid x = A \}}$. This is the set $\mathbf{\{ A \}}$ which is described in the conclusion.■

It is important to recognize the difference between $\mathbf{A}$ and $\mathbf{\{ A \}}$, particularly since it is very tempting and natural (but **dangerously incorrect!!**) to abbreviate the latter to $\mathbf{A}$. As noted near the bottom of page 4 in Halmos,

A box that contains a hat and nothing else is not the same thing as a hat.

The preceding result yields a simple example of a nonempty set.

Corollary 3. *There is a nonempty set* $\mathbf{A}$ *such that* $\mathbf{x \in A}$ *if and only if* $\mathbf{x = Ø}$.■

Since we are discussing results involving the empty set, this is a good time to mention one of its basic properties.

Proposition 4. *For every set* $\mathbf{A}$ *we have* $\mathbf{Ø \subset A}$.

Proof. This is similar to the last paragraph of the preceding argument. Since nothing belongs to $\mathbf{Ø}$ the statement **"**$\mathbf{(\forall x)\ x \in Ø \Rightarrow x \in A}$**"** is vacuously true.■

Sets defined by finite lists

We would like to elaborate upon the argument in Proposition 2 to show that for each finite list of sets $\mathbf{A_1, \dots , A_n}$ there is a set $\mathbf{\{ A_1, \dots , A_n \}}$ such that $\mathbf{B \in \{ A_1, \dots , A_n \}}$ if and only if $\mathbf{B = A_k}$ for some choice of $\mathbf{k}$. In order to keep the discussion simple, we shall initially limit ourselves to the case where $\mathbf{n} = 2$.

PAIRING PROPERTY. *If* $\mathbf{A}$ *and* $\mathbf{B}$ *are two sets, then there iexists a third set* $\mathbf{C}$ *such that* $\mathbf{A \subset C}$ *and* $\mathbf{B \subset C}$.

Proposition 5. *Suppose that* $\mathbf{x}$ *and* $\mathbf{y}$ *are distinct sets. Then there is a set* $\mathbf{\{ x, y \}}$ *(the* ***unordered pair****) such that* $\mathbf{z \in \{ x, y \}}$ *if and only if* $\mathbf{z = x}$ *or* $\mathbf{z = y}$.

Proof. By the Pairing Axiom there is a set $\mathbf{C}$ such that $\mathbf{\{ x \} \subset C}$ *and* $\mathbf{\{ y \} \subset C}$. Therefore by the Specification Axiom there is a set defined by the description $\mathbf{\{ z \in C \mid z = x}$ or $\mathbf{z = y \}}$. This is precisely the set described in the conclusion.■

In most situations that arise in mathematics, if we are given a finite list of sets $\mathbf{A_1, \dots , A_n}$ then the underlying assumptions will imply the existence of a set $\mathbf{C}$ such that $\mathbf{A_k \in C}$ for all $\mathbf{k}$, and in such cases there is a simple generalization of the previous result.

<u>Proposition 6.</u> *Suppose that* $\mathbf{A_1, \ldots, A_n}$ *are sets, and assume also that there is some set* $\mathbf{C}$ *such that* $\mathbf{A_k} \in \mathbf{C}$ *for all* $\mathbf{k}$. *. Then there exists a set* $\{\mathbf{A_1, \ldots, A_n}\}$ *such that we have* $\mathbf{B} \in \{\mathbf{A_1, \ldots, A_n}\}$ *if and only if* $\mathbf{B} = \mathbf{A_k}$ *for some* $\mathbf{k}$.

<u>Proof.</u> In this case the desired set is given by the following condition:

$$\{ \mathbf{x} \in \mathbf{C} \mid \mathbf{x} = \mathbf{A_k} \text{ for some } \mathbf{k} \}$$

Equivalently, the set is also given by the following description:

$$\{ \mathbf{x} \in \mathbf{C} \mid \mathbf{x} = \mathbf{A_1} \text{ or } \mathbf{x} = \mathbf{A_2} \text{ or } \ldots \text{ or } \mathbf{x} = \mathbf{A_n} \}$$

Either way we obtain the desired set.■

<u>Further examples.</u> The middle paragraph on page 10 of Halmos gives several examples of sets that can be constructed using the information about set theory that we have covered up to this point.